

AD-A259 318



DTIC
ELECTE
JAN 5 1993
S c D

Army Science Conference Proceedings

22-25 June 1992

Volume I

Principal Authors A through G

**This document has been approved for public
release and sale; its distribution is unlimited.**

92 12 22 135

92-32557



47385

Assistant Secretary of the Army
(Research, Development and Acquisition)
Department of the Army



DEPARTMENT OF THE ARMY
OFFICE OF THE ASSISTANT SECRETARY
WASHINGTON, DC 20310-0103



SARD-TR

30 JUL 1992

MEMORANDUM FOR SEE DISTRIBUTION

SUBJECT: Proceedings of the 1992 Army Science Conference

The 18th Army Science Conference was held at the Hyatt Hotel, Orlando, Florida, 22-25 June 1992. The conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army (DA). Additionally, it provided an opportunity for DA civilian and military scientists and engineers to present the results of their research and development efforts before a distinguished and critical audience.

These Proceedings of the 1992 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted.

Our purpose for soliciting these papers was to:

- a. Stimulate the involvement of scientific and engineering talent within the Department of the Army;
- b. Demonstrate Army competence in research and development;
- c. Provide a forum wherein Army personnel can demonstrate the full scope and depth of their current projects; and
- d. Promote the interchange of ideas among members of the Army scientific and engineering community as well as management and technical personnel from industry and other government agencies.

The information contained in these volumes will be of benefit to those who attended the conference and to others interested in Army research and development. It is requested that these Proceedings be placed in technical libraries where they will be available for reference.

George T. Singley III
Deputy Assistant Secretary
For Research and Technology

DISTRIBUTION:

Office of the Under Secretary of Defense for Research and Advanced Technology, Wash, DC 20310-3080

Office of the Assistant Secretary of the Army (RD&A), ATTN: SARD-TR, Wash, DC 20310-0103

HQDA, Assistant Chief of Staff for Intelligence, ATTN: DAMI-ZA, Wash, DC 20310

HQDA, Deputy Chief of Staff for Logistics, ATTN: DALO-ZX, Wash, DC 20310-0500

HQDA, Deputy Chief of Staff for Operations & Plans, ATTN: DAMO-FD, Wash, DC 20310-0400

HQDA, Director of Information Systems for C4, ATTN: SAIS-ZA, Wash, DC 20310-0400

HQDA, Deputy Chief of Staff for Personnel, ATTN: DAPE-ZA, Wash, DC 20310-0300

Office of the Chief of Engineers, ATTN: CERD-M, 20 Massachusetts Avenue, N.W., Wash, DC 20314-1000

Office of the Surgeon General, ATTN: DASG-ZA, 5109 Leesburg Pike, Falls Church, VA 22041-3258

COMMANDERS/DIRECTORS:

USA Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20814-2797

USA Operational Test & Evaluation Agency, 5600 Columbia Pike, Falls Church, VA 22041-5115

USA Information Systems Engineering Command, Ft. Huachuca, AZ 85613-5000

USA Strategic Defense Command, P. O. Box 15280, Arlington, VA 22215-0150

USA Materiel Command, ATTN: AMCLD, 5001 Eisenhower Avenue, Alexandria, VA 22333-0001

USA Materiel Systems Analysis Activity, ATTN: AMXS-D, APG, MD 21005-5071

USA Armament, Munitions & Chemical Command, ATTN: AMSMC-CG, Rock Island, IL 61299-6000

USA Armament RD&E Center, ATTN: SMCAR-TD, Picatinny Arsenal, NJ 07806-5000

Fire Support Armament Cen, ATTN: SMCAR-FS, PA, NJ 07806-5000

Armament Engineering Dir, ATTN: SMCAR-AE, PA, NJ 07806-5000

Close Combat Armament Cen, ATTN: SMCAR-CC, PA, NJ 07806-5000

Chemical RD&E Center, ATTN: SMCCR-TD, APG, MD 21010-5423

USA Aviation Systems CMD, ATTN: AMSAV-GTD, 4300 Goodfellow Blvd, St. Louis, MO 63120-1798

USARTA, Ames Research Center, ATTN: SAVRT-D, Moffett Field, CA 94035-1099

USARTA, Aeroflightdynamics Dir, ATTN: SAVRT-AF, Moffett Field, CA 94035

USARTA, Aviation Applied Technology Dir, ATTN: SAVRT-TY, Ft Eustis, VA 23604

USARTA, Propulsion Dir, ATTN: SAVRT-PN, Lewis Research Center, 21000 Brook Park Rd, Cleveland, OH 44135

USARTA, Aerostructures Dir, ATTN: SAVRT-SD, Hampton, VA 22665

Avionics R&D Activity, ATTN: SAVAA, Ft Monmouth, NJ 07703-5000

USA Communications-Electronics Cmd, ATTN: AMSEL-TDD, Ft Monmouth,
 NJ 07703-5001
 Center for Command, Control & Communications System, ATTN:
 AMSEL-RD-C3, Ft Monmouth, NJ 07703-5202
 Center for Night Vision & Electro-Optics, ATTN: AMSEL-RD-NV-D, Ft
 Belvoir, VA 22060-5677
 Electronics Warfare/RSTA Center, ATTN: AMSEL-RD-EW-D, Ft Monmouth,
 NJ 07703-5303
 Signals Warfare Lab, ATTN: AMSEL-RD-SW-D, Vint Hill Farms Station,
 Warrenton, VA 22186-5100
 USA Laboratory Command, ATTN: AMSLC-TD, 2800 Powder Mill Road,
 Adelphi, MD 29783-1145
 Atmospheric Sciences Lab, ATTN: SLCAS-DD, WSMR, NM 88002-5501
 Ballistic Research Lab, ATTN: SLCBR-OD, APG, MD 21005
 Electronics Technology & Devices Lab, ATTN: SLCET-D, Ft Monmouth,
 NJ 07703-5000
 Electronic Warfare Vulnerability Assessment Lab, ATTN: SLCEW-M-D,
 WSMR, NM 8802-5513
 Harry Diamond Labs, ATTN: SLCHD-D, Adelphi, MD 20783-1197
 Human Engineering Lab, ATTN: SLCHE-H, APG, MD 21005-5001
 Materials Technology Lab, ATTN: SLCMT-D, Watertown, MA 02172-0001
 Army Research Office, ATTN: SLCRO-D, RTP, NC 27709-2211
 Army Research, Development & Standardization Group (Europe) Box 65,
 FPO New York 09510
 USA Missile Command, ATTN: AMSMI-R, Redstone Arsenal, AL 35898-5240
 USA Tank-Automotive Cmd, ATTN: AMSTA-CG, Warren, MI 48397-5000
 Aberdeen Proving Ground, ATTN: STEAP-OC, APG, MD 21005-5001
 Dugway Proving Ground, ATTN: STEDP-CO, Dugway, UT 84022-5000
 Electronic Proving Ground, ATTN: STEEP-DO, Ft Huachuca, AZ
 85613-7110
 Aviation Development Test Activity, ATTN: STEBG-CO, Ft Rucker, AL
 36362-5276
 Combat Systems Test Act, ATTN: STECS-CO, APG, MD 21005-5059
 White Sands Missile Range, ATTN: STEWS-CG, WSMR, NM 88002-5000
 USA Aviation Systems Cmd, ATTN: AMSAV-G, 4300 Goodfellow Blvd, St.
 Louis, MO 63120-1798
 Belvoir RD&E Center, ATTN: STRBE-Z, Ft. Belvoir, VA 21005-5606
 Natick RD&E Center, ATTN: STRNC-Z, Natick, MA 01760-5000
 US Army Corps of Engineers
 Cold Regions Rsch & Eng Lab, ATTN: CRREL-TD, P. O. Box 282,
 Hanover, NH 03755-1290
 Construction Eng Rsch Lab, ATTN: CERL-Z, P. O. Box 4005,
 Champaign, IL 61820-135
 Engineer Topographic Labs, ATTN: ETL-TD, Ft Belvoir, VA
 22060-5546
 Waterways Experiment Station, ATTN: CEWES-IM-MI, P. O. Box 631,
 Vicksburg, MS 39180-0631
 USA Medical R&D Command, ATTN: SGRD-ZA, Ft Detrick, Frederick, MD
 21701-5012
 Aeromedical Rsch Lab, ATTN: SGRD-UAC, Ft Rucker, AL 36362-5292
 Inst of Dental Rsch, WRAMC, ATTN:SGRD-UDZ, Wash, DC 20307-5300
 Inst of Surgical Rsch, ATTN: SGRD-USZ, Ft Sam Houston, TX
 78234-6200

Letterman Army Inst of Rsch, ATTN: SGRD-ULZ, Presidio of San Francisco, CA 94129-6800
 Biomedical R&D Lab, ATTN: SGRD-UBZ, Frederick, MD 20701-5010
 Medical Rsch Inst of Chemical Defense, ATTN: SGRD-UV-ZA, APG, MD 21010-5425
 Medical Rsch Inst of Environmental Medicine, ATTN: SGRD-UEZ, Natick, MA 01760-5007
 Medical Rsch Inst of Infectious Diseases, ATTN: SGRD-UIZ-A, Ft Detrick, Frederick, MD 21701-5011
 Walter Reed Army Inst of Rsch, ATTN: SGRD-UWZ, Washington, DC 20307-5100
 USA Health Services Command, Ft. Sam Houston, TX 78234-61000
 USA Environmental Hygiene Agency, ATTN: HSHB-Z, APG, MD 21010-5422
 USA Research Institute for the Behavioral and Social Sciences, ATTN: PERI-ZT, 5001 Eisenhower Ave, Alexandria, VA 22333-5600
 ARI Field Unit, ATTN: PERI-IJ, POB 2086, Ft Benning, GA 31905
 ARI Field Unit, ATTN: PERI-SB, POB 6057, Ft Bliss, TX 79906
 ARI Field Unit, ATTN: PERI-SH, Ft Hood, TX 76544-5065
 ARI Field Unit, ATTN: PERI-IK, Steele Hall, Ft Knox, KY 40121
 ARI Field Unit, ATTN: PERI-SL, POB 290, Ft Leavenworth, KS 66027
 ARI Field Unit, ATTN: PERI-IR, Ft Rucker, AL 36362-5000
 ARI Field Unit, ATTN: PERI-IO, POB 5787, Presidio of Monterey, CA 93940
 USA Training and Doctrine Command, ATTN: ATDO-ZT, Ft Monroe, VA 23651-5000
 Combined Army Test Act, ATTN: ATCT-CG, Ft Hood, TX 76544-5065
COMMANDANTS:
 USA Air Defense Cen, ATTN: ATZC-CG, Ft Bliss, TX 79916-5000
 USA Armor Cen, ATTN: ATZK-CG, Ft Knox, KY 40121-5000
 USA Aviation Cen, ATTN: ATZQ-CG, Ft Rucker, AL 36362-5000
 USA Chemical Cen, ATTN: ATZN-CG, Ft McClellan, AL 36205-5000
 USA Combined Arms Cen, ATTN: ATZL-SC, Ft Leavenworth, KS 66027-5000
 USA Engineer Cen, ATTN: ATZA-CG, Ft Belvoir, VA 22060-5000
 USA Infantry Cen, ATTN: ATZB-CG, Ft Benning, GA 31905-5000
 USA Intelligence Cen, ATTN: ATSI-SA, Ft Hauchuca, AZ 85613-7000
 USA Ordnance Cen, ATTN: ATSL-CMT, APG, MD 21005-5201
 USA Signal Cen, ATTN: ATZH-CH, Ft Gordon, GA 30905-5000
 USA Field Artillery Center, ATTN: ATZR-C, Ft Sill, OK 73503

SUPERINTENDENT:

US Military Academy, ATTN: Technical Library, West Point, NY 10996

COPIES FURNISHED:

Defense Advanced Research Projects Agency, 1400 Wilson Blvd, Arlington, VA 22209-2308
 Defense Logistics Agency, Cameron Station, Alexandria, VA 22304-6183
 Defense Technical Information Center, Cameron Station, Alexandria, VA 22304-6145
 Lawrence Livermore National Lab, ATTN: L-191, POB 808, Livermore, CA 94550
 Los Alamos National Lab, ATTN: Dir for Energy, Rsch & Tech, Los Alamos, NM 87545

NASA HQS, Suite 4237, 400 Maryland Ave, SW, Wash, DC 20546
National Science Foundation, 1800 G Street, NW, Wash, DC 20550
United Nations Library, ATTN: Acquisition Section, Room L-138A,
New York, NY 10017

US NAVY

Naval Air Systems Command, Code AIR-03-D, 1411 Jefferson Davis
Hwy, Arlington, VA 22202-3000
Naval Rsch Lab, ATTN: Dir of Rsch, Washington, DC 20375
Office of Naval Rsch, Code 10, 800 North Quincy St, Arlington, VA
22217-5000

HQ US Marine Corps, Code RD-1, Washington, DC 29380-0001

US AIR FORCE

Air Force Systems Command, ATTN: Technical Director, Andrews AFB,
Washington, DC 20334-5000
Air Force Office of Scientific Research/XOT, ATTN: Distribution,
Bolling AFB, Washington, DC 20332-6448

DTIC QUALITY INSPECTED 1

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

PROCEEDINGS
OF THE
1992 ARMY SCIENCE CONFERENCE

ORLANDO, FLORIDA
22 - 25 JUNE 1992

VOLUME I
PRINCIPAL AUTHORS A THROUGH G

TABLE OF CONTENTS

PROCEEDINGS OF THE 1992 ARMY SCIENCE CONFERENCE

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Alabran, David M.	See Willingham, Reginald	III	369
Alster, Jack	See Nicolich, Steven	III	1
Andersen, Steven L.	See Martin, Rodger K.	II	301
Atha, Larry C.	Advanced Composite Interceptor Kill Vehicle Structures	I	5
Azrin, Morris	See Beatty, John H.	I	51
Bahder, Thomas B.	Analytic Dispersion Relations Near the Gamma Point in Strained Zincblende Semiconductors	I	19
Baker, Paul K.	Lightweight Exo-Atmospheric Projectile (LEAP)	I	25
Balekdjian, Armand	See Harvey, James F.	II	45
Barrows, Austin W.	Development of Laser-Based Ignition Systems for Large Caliber Guns	I	37
Beatty, John H.	Correlation of Ballistic Performance to Shear Instability Studies in High Strength Steels	I	51
Beaudry, William T.	See Wagner, George W.	III	247
Behrens, Richard, Jr.	Thermal Decomposition of RDX and HMX in the Condensed Phase: Isotope Scrambling and Deuterium Isotope Effect Studies	I	65
Bennett, Kelly W.	See Brody, Philip S.	I	175
Berger, Henry	A New Understanding of Power Transmission and Reflection in Remote and Near-In Sensing	I	83
Beyer, Richard A.	Small Scale Experiments and Modeling of Laser Ignition of Gun Propellants	I	97

	<u>VOL</u>	<u>PAGE</u>
, Austin W.	I	37
Robert H.	III	35
o, Janine	I	471
Richard L.	III	221
, Wendell R.	III	275
Increasing Breakdown Polymer Film	I	111
, Michael J.	II	353
of the Mine ehicle Using Advanced al Methods	I	119
ham, Reginald	III	369
, Michael J.	II	353
Arpad	II	159
echnologies for the Twenty-First Century	I	1
Robert H.	III	35
Characteristics of enerated Charge in nsulator Buried	I	135
r, Michael E.	I	319
James F.	II	45
reba, Steven	III	119
ary S.	III	231
ctronic Detonator Live-Fire Missile	I	151

<u>THOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Brimfield, A. A.	Catalytic Antibodies as Biological Scavengers for Organophosphorus Poisons	I	167
Brody, Philip S.	Photovoltaic Effect in Thin Ferroelectric Films for Nonvolatile Memory	I	175
Broomfield, C. A.	See Brimfield, A. A.	I	167
Brown, Steven D.	See Harper, Alice M.	II	31
Brown, C. David	Scene Segmentation and Definition for Autonomous Robotic Navigation Using Structured Light Processing	I	189
Bryant, Garnett W.	See Tober, Richard L.	III	221
Bryant, Garnett W.	See Tobin, Mary S.	III	231
Bryzik, Walter	Advanced Ground Propulsion System Phenomena	I	205
Bulusu, Suryanarayana	See Behrens, Richard Jr.	I	65
Burchanowski, John	See Guadagno, Janine	I	471
Barrillo, Alejandro	See Horner, David A.	II	109
Cartland, Harry E.	Time-Resolved Infrared Diode Laser Spectroscopy: A State Specific Probe of Atom-Molecule Reaction Dynamics	I	221
Casas, L.	See Lee, Howard S.	II	253
Chambers, James P.	See Valdes, James J.	III	239
Chang, Wayne H.	See Cooke, P. W.	I	291
Chang, Wayne H.	See Choi, K. K.	I	253
Chang, Wayne H.	See Lux, Robert A.	II	295
Chang, Wayne H.	See Shen, Hongen	III	133

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Chang, Wayne H.	See Dutta, Mitra	I	399
Checkai, Ronald	See Wentzel, Randall	III	345
Chin, Ernest S. C.	Dynamic Impact Response of Titanium Aluminide Composites	I	237
Choi, K. K.	High Sensitivity Infrared Hot- Electron Transistors	I	253
Christensen, Charles	Optical Mapping of HgCdTe Detector Electronic Properties	I	267
Chu, Shih C.	A More Rational and Rigorous Analysis of In-Bore Projectile	I	277
Clarke, Frederick W.	See Christensen, Charles	I	267
Clayton, Michael	See Sjogren, Maria H.	III	161
Cohen, Arthur	See Beyer, Richard A.	I	97
Cohen, Arthur	See Barrows, Austin W.	I	37
Cole, Melanie W.	See McLane, G.F.	II	339
Cole, Melanie W.	See Lee, Howard S.	II	253
Cole, Melanie W.	See Harvey, James F.	II	45
Colston, James T.	See Valdes, James J.	III	239
Condie, Richard M.	See Sjogren, Maria H.	III	161
Cooke, P. W.	Strained Quantum Well Field Effect Lasers for Optoelectronic Integration	I	291
Cornell, John H.	See Willingham, Reginald	III	369
Cotariu, Steven S.	Transient Energy Transfer in Wide Gap Semiconductors	I	305
Crenshaw, Michael E.	New Optical Switch Stemming from Dipole-dipole Interactions in Dense Media	I	319
Crow, Samuel B.	See Watkins, Wendell R.	III	275

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Cullen, John W.	See Willingham, Reginald	III	369
Danielson, Eugene	See Bryzik, Walter	I	205
Daschka, Paul	Ceramic Matrix Fiber Reinforced Composites	I	333
Davis, David	See Reid, Robert H.	III	35
DeAnni, Anthony	See Lux, Robert A.	II	295
DeAnni, Anthony	See Lee, Howard S.	II	253
DeCamp, Brian C.	MAXIMIZE Automatic Target Recognition Algorithm Development Environment	I	349
DeFrank, Joseph J.	See Harvey, Steven P.	II	59
DiDomenico, Leo	See Mizan, Muhammad	II	397
Dillon, Tom M.	A Framework for the Environmental Risk Assessment of Contaminated Sediments at DoD Installations	I	357
Don, R. C.	Ballistic Performance--Material Property Relationships of Glass Armor Versus Shaped Charge Jets	I	373
Doughty, David T.	See Morrison, Robert W.	II	409
Dowding, Robert J.	Current Status of Tungsten Alloy Research	I	383
Dubey, Madan	See Harvey, James F.	II	45
Dudley, M.	See Fazi, C.	I	421
Dutta, Mitra	See Harvey, James F.	II	45
Dutta, Mitra	A Novel High Contrast Optical Modulator in a GaAs-AlAs Structure	I	399
Dutta, Mitra	See Shen, Hongen	III	133

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Eckart, D.W.	See McLane, G.F.	II	339
Ehlen, Judy	See Hansen, John V.E.	II	1
Eicke, John S.	See Wiles, George C.	III	357
Eidelman, Shmuel	See Bishop, Steven S.	I	119
Eldefrawi, Moyhee D.	See Thompson, Roy G.	III	205
Ennis, Douglas H.	See Atha, Larry C.	I	5
Esch, Frederick H.	Imagery Exploitation System/Balanced Technology Initiative	I	407
Evaldsson, P. A.	See Cooke, P. W.	I	291
Evans, Timothy D.	See Hansen, John V.E.	II	1
Ezzell, John W.	See Friedlander, Arthur	I	445
Fazi, C.	Observation of rf-Induced Failures in Silicon Bipolar Junctions Using Synchrotron X-Ray Topographic Imaging	I	421
Ferrando, Ronald E.	See Schuschereba, Steven	III	119
Fisher, Jamie M.	Characterizing the Response of Solid Rocket Motors to Fragment Impact	I	429
Forch, Brad E.	See Barrows, Austin W.	I	37
Franz, David R.	See Sjogren, Maria H.	III	161
Friedlander, Arthur	Post-exposure Prophylaxis Against Experimental Inhalation Anthrax	I	445
Gatto, Lawrence	See Melloh, Rae	II	381
Gentry, Mary K.	See Lebeda, Frank J.	II	237
Gillespie, James B.	Battlefield Environmental Effects on Ultraviolet Atmospheric Sensing Propagation	I	459

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Gordon-Smith, D.	See Fazi, C.	I	421
Gorman, Francis J.	See Soicher, Haim	III	193
Grisham, John A.	See Christensen, Charles	I	267
Grogl, Max	See Martin, Rodger K.	II	301
Guadagno, Janine	Landmine Detection by Backscatter Radiation Radiography	I	471
Gupta, N.	Optical Neural Network Device Architecture	I	485
Hack, Dallas C.	See Lebeda, Frank J.	II	237
Han, W.Y.	See Lee, Howard S.	II	253
Hansen, John V.E.	Mine Detection in Dry Soils Using Radar	II	1
Haraburda, Scott S.	Developmental Research for Designing a Microwave Electrothermal Thruster	II	15
Hardaway, Mike	See Welsh, James P.	III	335
Harnden, Phillip G.	See Shires, Dale R.	III	147
Harper, Alice M.	Artificial Intelligence for Chemical and Biological Detectors	II	31
Harvey, James F.	See Lux, Robert A.	II	295
Harvey, James F.	Physics and Device Applications of Silicon Microclusters	II	45
Harvey, Steven P.	Biodegradation of Chemical Warfare Agents: Demilitarization Applications	II	59
Hayes, Pamela G.	Vulnerability of Protective Structures to Artillery Fire	II	71
Hepfinger, Lisa B.	Laser-Induced Luminescence: Soldier Vulnerability	II	83

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Hevenor, Richard A.	See Hansen, John V.E.	II	1
Higgins, Thomas P.	See Lux, Robert A.	II	295
Hock, Vincent F.	See Ruzga, Richard J.	III	63
Holly, Frank F.	See Shires, Dale R.	III	147
Hoock, Donald W.	An Approach to Mitigating Atmospheric Effects on Image- Based Pattern Recognition by Neural Networks	II	95
Horner, David A.	Optimization of Particle Model for Large-Scale Computing	II	109
Howe, Gerald B.	See Friedlander, Arthur	I	445
Huff, William L.	See Hayes, Pamela G.	II	71
Hunt, Robert E.	Host Respiratory Protection for Lethal Staphylococcal Enterotoxin B	II	119
Hursh, Steven R.	Modeling Human Performance to Predict Unit Effectiveness	II	129
Ivins, Bruce E.	See Friedlander, Arthur	I	445
Jasper, L.	See Kim, A.	II	197
Jasper, L.	See Kim, A.	II	185
Jett, Marti	See Hunt, Robert E.	II	119
Johnson, Anthony J.	See Hunt, Robert E.	II	119
Johnson, Daniel L.	See Patterson, James H.	III	25
Johnson, John C.	High Bandwidth, Wide Field of View Lasercom Demonstration	II	145
Johnson, M.A.	See Meisel, L.V.	II	365
Jones, K.A.	See Lee, Howard S.	II	253
Juhasz, Arpad	Solid Propellant Electrothermal Gun Propulsion	II	159

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Kaplan, Zvi	See Juhasz, Arpad	II	159
Katechis, James C.	Dormancy for the Ground Based Interceptor	II	173
Kim, Hie-Joon	See Ross, Edward	III	49
Kim, A.	Monolithic, Photoconductive Impulse-Generating Device	II	197
Kim, A.	High-Power Impulse Generation Using a Photoconductively Switched Radial Transmission Line	II	185
King, Wendell C.	See Witten, Alan J.	III	383
Klohn, Ken	See Mizan, Muhammad	II	397
Klopacic, J. Terrence	See Walbert, James N.	III	261
Kodak, James A.	See Reid, Robert H.	III	35
Komisar, Jack L.	See Hunt, Robert E.	II	119
Koscica, Thomas E.	See Lux, Robert A.	II	295
Kumar, Ashok	See Daschka, Paul	I	333
Kumar, Pramod	See Valdes, James J.	III	239
Kustin, Kenneth	See Ross, Edward	III	49
Kyle, Dennis E.	See Martin, Rodger K.	II	301
Ladas, Andrew P.	See Wiles, George C.	III	357
LaiHing, Kenneth	See McDonald, Joseph K.	II	327
Lane, Gerald R.	Establishing Integratged Two-Man Crew Station (ITCS) Requirements	II	209
Lanzerotti, Yvonne	Power Spectral Characterization of Fracture Surfaces of TNT, Composition B, and Octol	II	223
Lareau, Richard T.	See McLane, G.F.	II	339

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Lareau, Richard T.	See Harvey, James F.	II	45
Lareau, Richard T.	See Lee, Howard S.	II	253
Lareau, Richard T.	See Schauer, Stephen N.	III	105
LaSala, John E.	See Cotariu, Steven S.	I	305
Lavene, Bernard	See Binder, Michael	I	111
Lawrence, Wade B.	See Friedlander, Arthur	I	445
Leavitt, Richard P.	See Tobin, Mary S.	III	231
Leavitt, Richard P.	See Gupta, N.	I	485
Lebeda, Frank J.	Thermodynamic Analyses of Transmembrane Channel Formation by Botulinum Toxin	II	237
Lee, Howard S.	See Schauer, Stephen N.	III	105
Lee, Howard S.	See McLane, G.F.	II	339
Lee, Howard S.	Ohmic Contacts to Heavily Carbon-Doped p ⁺ -GaAs Using Ti/Si/Pd	II	253
Lemire, George W.	See Sausa, Rosario C.	III	91
Lenz, D. E.	See Brimfield, A. A.	I	167
Lepore, A.	See McLane, G.F.	II	339
Li, W. Q.	See Tober, Richard L.	III	221
Lieb, Robert J.	A Small Angle Neutron and X-Ray Scattering Study of the Onset and Nature of Fracture of Uniaxially Compressed Gun Propellants	II	267
Liston, John E.	Exoatmospheric Intercept: GBI Program and the ERIS FTV Flight Tests	II	281
Lottati, Isaac	See Bishop, Steven S.	I	119

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Lowe, John R.	See Friedlander, Arthur	I	445
Lu, Y.	See Lee, Howard S.	II	253
Lukasavage, William	See Nicolich, Steven	III	1
Lukaszek, Ted	See Mizan, Muhammad	II	397
Lund, David J.	See Schuschereba, Steven	III	119
Lux, Robert A.	Optical Control of Resonant Tunnel Diode Devices	II	295
Lux, Robert A.	See Harvey, James F.	II	45
Mammone, Robert J.	See Binder, Michael	I	111
Martin, Rodger K.	<i>Plasmodium falciparum</i> : Characterization of Multiple Drug Resistance Phenotypes by Reversal Modulators	II	301
Matthew, Candace B.	An Animal Model of Drug-Induced Thermoregulatory and Endurance Decrements	II	315
Maxwell, D. M.	See Brimfield, A. A.	I	167
Mays, Brian T.	See Wiles, George C.	III	357
McDonald, Joseph K.	Nonlinear Optical Properties of Layered Composites	II	327
McKee, Kelly T.	See Sjogren, Maria H.	III	161
McLane, G.F.	Magnetron Ion Etching for GaAs Device Processing	II	339
McNally, Richard E.	See Hursh, Steven R.	II	129
McQuaid, Michael J.	Imaging and Analysis of Combusting Liquid Propellant Sprays	II	353
McQueen, Charles E.	See Reid, Robert H.	III	35
Meisel, L.V.	Multifractal Analysis of Chaotic Point Sets	II	365

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Melloh, Rae	River and Lake Ice Conditions as Interpreted from Microwave Imagery	II	381
Melnik, D.	See Juhasz, Arpad	II	159
Menking, Darrel E.	See Thompson, Roy G.	III	205
Meyyappan, M.	See McLane, G.F.	II	339
Mikesell, Perry	See Friedlander, Arthur	I	445
Milhous, Wilbur K.	See Martin, Rodger K.	II	301
Mizan, Muhammad	Microwave Solid-State Transmitter for Miniature Multiband Beacon Transponder	II	397
Miziolek, Andrzej W.	See Barrows, Austin W.	I	37
Miziolek, Andrzej W.	See Sausa, Rosario C.	III	91
Morrison, Robert W.	Chromium-free Carbon, a New Adsorbent for Chemical Warfare Agent Filters	II	409
Morton, David C.	See Harvey, James F.	II	45
Mozo, Ben T.	See Patterson, James H.	III	25
Nakashima, Masato	See Willingham, Reginald	III	369
Namaroff, M.	See McLane, G.F.	II	339
Nelson, William R.	See Patterson, James H.	III	25
Newberry, Joyce E.	See Barrows, Austin W.	I	37
Newman, Peter G.	See Dutta, Mitra	I	399
Newman, Peter G.	See Lux, Robert A.	II	295
Nickolaisen, Scott	See Cartland, Harry E.	I	221
Nicolich, Steven	Quantitative Synthesis of HMX via the GARDEC HMX Process	III	1

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Norvelle, F. Raye	Using Iterative Orthophoto Refinements to Correct Digital Elevation Models	III	13
Nuzum, Edwin O.	See Martin, Rodger K.	II	301
Overholt, James L.	See Wehage, Roger A.	III	307
Palacios, Fernando	See Watkins, Wendell R.	III	275
Pamulapati, Jagadeesh	See Dutta, Mitra	I	399
Pamulapati, Jagadeesh	See Harvey, James F.	II	45
Paolella, Arthur C.	See Lux, Robert A.	II	295
Parmelee, Robert	See Wentzel, Randall	III	345
Patterson, James H.	New Human Exposure Limits for Freefield Artillery Blast Overpressure	III	25
Pearson, Earl F.	See McDonald, Joseph K.	II	327
Pennise, Christine	See Boesch, H. Edwin, Jr.	I	135
Peters, Clarence J.	See Sjogren, Maria H.	III	161
Peters, John F.	See Horner, David A.	II	109
Petreanu, John	See Daschka, Paul	I	333
Pham, John	See Gupta, N.	I	485
Pham, John	See Tobin, Mary S.	III	231
Phillips, Yancy Y.	See Patterson, James H.	III	25
Piekarz, Richard C.	See Harvey, James F.	II	45
Pinto, James J.	See Lanzerotti, Yvonne	II	223
Pitt, Louise	See Hunt, Robert E.	II	119
Pitt, Margaret L.M.	See Friedlander, Arthur	I	445
Quong, Julie A.	See Schuschereba, Steven	III	119

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Rael, Eppie D.	See Valdes, James J.	III	239
Rapacki, E. J., Jr.	See Don, R. C.	I	373
Reid, Michael	See Bryzik, Walter	I	205
Reid, Robert H.	Oral, Microencapsulated, CFA/II Vaccine Against <u>E. coli</u> Diarrheal Disease: Preclinical Evaluation	III	35
Remy, David E.	See Willingham, Reginald	III	369
Richmond, Donald R.	See Patterson, James H.	III	25
Ripple, Gary	See Patterson, James H.	III	25
Roach, Joseph F.	See Willingham, Reginald	III	369
Roberson, William E.	See Liston, John E.	II	281
Rod, B. J.	See Brody, Philip S.	I	175
Rogers, Kim R.	See Thompson, Roy G.	III	205
Rondeau, Ernest	See Binder, Michael	I	111
Rose, Kenneth J.	See Friedlander, Arthur	I	445
Rosen, David L.	See Gillespie, James B.	I	459
Ross, Edward	Kinetic Model for Predicting Bacterial Destruction from Intrinsic Chemical Marker Formation in Thermally Processed Foods	III	49
Ruble, David	See Hunt, Robert E.	II	119
Ruzga, Richard J.	See Daschka, Paul	I	333
Ruzga, Richard J.	The Infiltration of Cotton Fibers with Sol-Gel Solutions, Forming Oxide and Non-Oxide Ceramics	III	63

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Sadoff, Gerald C.	Cloning, Expression, Production and X-ray Crystallographic Structure of Acetylcholinesterase	III	79
Saphier, D.	See Juhasz, Arpad	II	159
Sasserath, J.	See McLane, G.F.	II	339
Sau, Keya	See Reid, Robert H.	III	35
Sausa, Rosario C.	Laser-Based Sensitive Detection of Trace Atmospheric Vapors of Military Interest	III	91
Scalora, Michael	See Crenshaw, Michael E.	I	319
Schauer, Stephen N.	Accurate Measurement of Metal Diffusion in Electronic Materials Using Secondary Ion Mass Spectrometry and Scanning Auger Microscopy	III	105
Schiff, Lewis B.	See Weinacht, Paul	III	321
Schuschereba, Steven	Basic Fibroblast Growth Factor Accelerates Repair and Prevents Degeneration in Retinal Laser Lesions	III	119
Schwarz, Ernest	See Bryzik, Walter	I	205
Sergi, Sergio A.	See Bishop, Steven S.	I	119
Shafferman, Avigdor	See Sadoff, Jerald C.	III	79
Shattuck, Edgar	See Ross, Edward	III	49
Shen, Hongen	Non Contact, Non Destructive Optical Method for Measuring Electric Fields in GaAs, AlGaAs and InGaAs Heterostructures	III	133
Shen, Hongen	See Harvey, James F.	II	45
Shen, Hongen	See Dutta, Mitra	I	399

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Shires, Dale R.	High-Ratio Bandwidth Reduction of Video Imagery for Teleoperation	III	147
Silman, Israel	See Sadoff, Jerald C.	III	79
Simeonsson, Josef B.	See Sausa, Rosario C.	III	91
Simonis, G. J.	See Gupta, N.	I	485
Sjogren, Maria H.	Safety and Kinetics of a Despeciated Equine F(ab) ₂ Heptavalent Botulinum Antitoxin in Volunteers	III	161
Smith, Doran D.	Scaling Properties of the Magnetoexciton Problem	III	169
Smith, Leonard A.	Cloning, Characterization, and Expression of Animal Toxin Genes for Vaccine Development	III	179
Smith, Stan	See Juhasz, Arpad	II	159
Soicher, Haim	Regression/Correlation Analysis of Ionospheric Parameters at Midlatitudes	III	193
Soreq, Hermona	See Sadoff, Jerald C.	III	79
Starks, Michael W.	See Walbert, James N.	III	261
Stead, M. R.	See Gupta, N.	I	485
Sturek, Walter B.	See Weinacht, Paul	III	321
Sturzebecher, Dana	See Lux, Robert A.	II	295
Sussman, Joel	See Sadoff, Jerald C.	III	79
Tanton, George A.	See Christensen, Charles	I	267
Taub, Irwin	See Ross, Edward	III	49
Taylor, Thomas L.	See Boesch, H. Edwin, Jr.	I	135
Taylor, G. W.	See Cooke, P. W.	I	291

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Taysing-Lara, M.	See Choi, K. K.	I	253
Thompson, R.J.	See Lee, Howard S.	II	253
Thompson, Roy G.	Chem-Biodetection: Potentiometric and Fiber Optic Biosensor Test Beds	III	205
Thomson, David J.	See Lanzerotti, Yvonne	II	223
Thornton, Clarence	See Harvey, James F.	II	45
Tober, Richard L.	Electric Field Effects on the Optical Properties of Layered Semiconductor Structures	III	221
Tobin, Mary S.	Simultaneous Electron and Hole Tunneling in Coupled Quantum Wells	III	231
Tressler, L.	See Don, R. C.	I	373
Trevino, Samuel F.	See Lieb, Robert J.	II	267
Tseng, Lee-Ying	See Reid, Robert H.	III	35
Tseng, Jeenan	See Hunt, Robert E.	II	119
Tsu, Raphael	See Harvey, James F.	II	45
Ursic, James R.	See Witten, Alan J.	III	383
Valdes, James J.	Detection of Picogram Quantities of Botulinum Toxin-B (BoTX) Using the Light Addressable Potentiometric Sensor	III	239
Valdes, James J.	See Thompson, Roy G.	III	205
Vanderbeek, Richard	See Harper, Alice M.	II	31
Vargas, Joseph A.	See Schuschereba, Steven	III	119
Veney, David W.	See Cartland, Harry E.	I	221
Wade, William L.	See Binder, Michael	I	111

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Wagner, George W.	Solid-State ³¹ P MAS NMR Study of G-Agent Simulants Adsorbed on Synthetic Resin Catalysts	III	247
Walbert, James N.	A Taxonomy for the Vulnerability/Lethality Analysis Process	III	261
Walker, John J.	See Katechis, James C.	II	173
Wang, S.	See Fazi, C.	I	421
Ward, J. Richard	See Wagner, George W.	III	247
Watkins, Wendell R.	Camouflage Flicker: A Highly Observable Low Observable	III	275
Watson, Jerry L.	Propellant Response to Shaped Charge Jet Impacts	III	291
Wehage, Roger A.	New Directions in Computational Dynamics at TACOM	III	307
Weinacht, Paul	Navier-Stokes Predictions of Pitch Damping for Axisymmetric Shell Using Steady Coning Motion	III	321
Weiner, M.	See Kim, A.	II	197
Weiner, M.	See Kim, A.	II	185
Welkos, Susan L.	See Friedlander, Arthur	I	445
Welsh, James P.	Review of Environmental Research Specific to Smart Weapons Operability Enhancement for the Battlefield Environment	III	335
Wentsel, Randall	Novel Techniques to Determine Ecological Effects at U.S. Army Sites	III	345
West, Wade	See Welsh, James P.	III	335
Wiles, George C.	Projectile Tracking Device Using GPS	III	357
Wilhelmsen, Catherine	See Hunt, Robert E.	II	119

<u>AUTHOR</u>	<u>TITLE</u>	<u>VOL</u>	<u>PAGE</u>
Wilhelmsen, Catherine	See Reid, Robert H.	III	35
Willingham, Reginald	Synthesis of Metallo-Tetrabenzoporphyrins Possessing High Third-Order Optical Nonlinearity for Military Laser Eye Protection	III	369
Witten, Alan J.	High Resolution Image Processing of Geophysical Data with Diffraction Tomography	III	383
Wittig, Curt	See Cartland, Harry E.	I	221
Wolfe, Allan	See Lanzerotti, Yvonne	II	223
Woolsey, Patrick	See Chin, Ernest S. C.	I	237
Worsham, Patricia L.	See Friedlander, Arthur	I	445
Yalamanchili, Rao	Aerodynamic Heating of Unsteady Hypervelocity Projectile	III	399
Yang, Janet M.	See Schuschereba, Steven	III	119
Yang, L.W.	See Lee, Howard S.	II	253
Youmans, R.	See Kim, A.	II	197
Youmans, R.	See Kim, A.	II	185
Zeto, R.	See Kim, A.	II	197
Zeto, R.	See Kim, A.	II	185
Zhou, Weimin	See Harvey, James F.	II	45
Zukas, Walter X.	Chemistry of Adhesives at Surfaces	III	417

Board on Army Science and Technology (BAST)

STAR 21
Strategic Technologies for the Army of the Twenty-First Century

Committee on Strategic Technologies for the Army (STAR)
National Academy of Sciences/BAST
Washington, D.C.

In the year 2020, the technical capabilities of the soldiers and systems that the United States Army puts into the field will result from the interplay of two kinds of processes. On one side are the processes of scientific research and technological innovation that determine what is *possible*. On the other are a wide array of factors that affect what is *needed*: requirements driven by the Army's mission, by the threats it faces, by political and economic changes at home and abroad, and by changes in the technology accessible to potential opponents. In the short run, the processes of *technology push* or those of *requirements pull* may dominate; in the longer term, factors of both kinds will influence the outcome.

This interplay of technology push and requirements pull has been a central theme in the reports prepared by the scores of experts who volunteered to participate in the National Research Council's Committee on Strategic Technologies for the Army (STAR). In requesting the study, the Assistant Secretary of the Army for Research, Development and Acquisition saw its purpose as assisting the Army in improving the introduction of advanced technologies into Army weapons, equipment, and doctrine. The time period addressed by the study was to extend at least 30 years into the future.

To conduct the study, the National Research Council organized three major subcommittees; Science and Technology, Integration, and Technology Management and Development Planning. Near the end of the study, a special committee, with representatives from these previously constituted subcommittees, was established to produce the STAR main report. The main report is available through the National Academy Press. A core briefing team was assembled in December 1991 and is led by Dr. Walter B. LaBerge, chairman of the special committee. He is supported by Dr. Joseph F. Soukup and Dr. John D. Venables, representatives of the STAR Science and Technology Subcommittee and Mr. David C. Hardison, the Systems Capabilities Subcommittee chairman. A short vita for each briefer follows.

Board on Army Science and Technology (BAST)

Briefer's Biographies

DR. WALTER B. LABERGE is presently Holder, Acquisition Policy Chair at the Defense Systems Management College, Ft. Belvoir, Virginia. He was appointed Assistant Secretary of the Air Force, Research & Development, and served until 1975 when he became Assistant Secretary General of NATO in Brussels, Belgium. In 1976, he was appointed Under Secretary of the Army, a post he held until 1979. He was one of the principal inventors of the Sidewinder air-to-air missile and led a team which designed the NASA Houston Mission Control Center for Apollo. He joined Lockheed Missiles & Space Company in July 1981 as Executive Assistant to the President, was appointed Vice President Planning and Technology in February 1982, and held that post until he was named to head the Research and Development Division. He also served as the principal deputy to Dr. William Perry, the former Under Secretary of Defense for Research and Engineering. Dr. LaBerge earned B.N.S., B.S., and Ph.D. degrees from the University of Notre Dame.

Dr. LaBerge has received numerous medals of commendation from the U.S. Air Force, U.S. Navy, U.S. Army, and the Department of Defense. He was elected to the National Academy of Engineering of the National Research Council (NRC) in 1987 and has served as chairman or member of several NRC committees and panels including his appointment as the chairman of the Study Committee on STAR. He is also a member of the Army Science Board.

MR. DAVID C. HARDISON presently supports the Department of Defense and several industrial firms as a consultant specializing in defense operations analysis. He has served as Director, U.S. Army Concepts Analysis Agency and as Deputy Under Secretary of Defense (Tactical Warfare Programs) from 1980 to 1982. He was Deputy Under Secretary, Army (Operations Research) from 1975 to 1980 and became the Senior U.S. representative to the NATO Patriot Acquisition Program. From 1952 to 1975 he held various positions as Analysis Advisor to the Deputy Chief of Staff for Research, Development and Acquisition, Scientific Advisor to the Commanding General, U.S. Army Combat Developments Command, and as Chief, Armor/Anti-Armor Systems, U.S. Army Ballistic Research Laboratories.

He earned an M.I.A degree from George Washington University, an M.A. degree from Duke University, a B.A. from Atlantic Christian College and a degree from the National War College. Mr. Hardison has served on several government advisory panels and is currently the Chairman of the National Research Council's (Board on Army Science and Technology) STAR Systems Capabilities Subcommittee and the Support Systems Panel.

Board on Army Science and Technology (BAST)

DR. JOSEPH F. SOUKUP is a Corporate Vice President of Science Applications International Corporation (SAIC) and manages the Life Sciences and the Chemical Demilitarization Divisions. He also serves as an independent advisor to the U.S. intelligence community in the areas of chemical, toxin, and biological threat analysis and biotechnology. He has continuously supported the Army Medical Research and Development Command in the areas of medical countermeasures to these toxin agents for the past nine years. Dr. Soukup has authored over 50 papers in neuroscience, classified technical reports on chemical and toxin agents and biotechnology, and was the senior manager of a current textbook on Medical Chemical Defense against mustard agents.

Dr. Soukup earned a Ph.D. in neuropharmacology from Duke University and conducted post doctoral research at Duke under a peer-competed National Institutes of Health grant. Prior to entering the life sciences field, he trained and worked for General Electric Co. in electronics engineering. Dr. Soukup is a member of AAAS, ADPA, and AUSA. He is currently serving as a member of the STAR Science and Technology Subcommittee as the Vice Lead Expert of the Biotechnology and Biochemistry Technology Group.

DR. JOHN D. VENABLES is presently a consultant and retired from Martin Marietta Laboratories where he held the position of Corporate Scientist. He was actively involved in many aerospace programs, including the Shuttle External Tank and the Viking Mars Lander. Work performed by Dr. Venables led to the discovery of several important phenomena, including ordered structures in the transition metal carbides, and has provided new insights into adhesive bonding mechanisms and factors controlling the durability of adhesive joints.

Dr. Venables has published over 60 technical papers, holds four patents, and has lectured at numerous conferences and universities. He is on the Editorial Advisory Board of the Journal of Adhesion Science and Technology, and M.I.T.'s *Materials & Processing Report*. He is a member of the American Physical Society, American Association for the Advancement of Science, and several other professional organizations. He has served on various National Academy of Sciences study committees and is presently a member of the STAR Science and Technology Subcommittee as Lead Expert of the Advanced Materials Technology Group.

ATHA, ENNIS

Advanced Composite Interceptor Kill Vehicle Structures

Dr. Larry C. Atha
U. S. Army Strategic Defense Command
Atten: CSSD-SL-K
Huntsville, Alabama 35807-3801

Mr. Douglas H. Ennis*
U. S. Army Strategic Defense Command
Atten: CSSD-SL-K
Huntsville, Alabama 35807-3801

(U) In 1987, following a late 1986 review of critical issues and the identification of resolution path options, the U. S. Army Strategic Defense Command (USASDC) initiated a development and evaluation sequence designed to determine if advanced composites could be applied to the main structural components of strategic interceptors. The program was constructed as a coordinated effort with the U. S. Army Materials Technology Laboratory (USAMTL) and focused upon practical application, not a simple demonstration that these materials could serve the same purpose as conventional materials. Responsibility for the execution of this work was assigned to the Key Technologies Division (KT) of the Survivability, Vulnerability, and Key Technologies Directorate (SLKT). KT incorporated the activity into its Materials and Structures Program (M&S). A three-fold approach was selected to meet the need for parallel development of those aspects of advanced composites which were either considered clearly inadequate, or questionable, by the user Project Offices:

- 1) Selection of the optimum composite for a given application from both the standpoint of performance and cost – with further material development if required and if justified in terms of the near-term potential of that material¹;
- 2) Parallel development of improved bonding and joining techniques to ensure that, to the maximum extent possible, the native properties of a selected material were not degraded when the material was used in a realistic assembly²;

- 3) Evaluation of the final assembly against its conventional counterpart using identical standards and, to the maximum extent consistent with valid test methodology, identical procedures³.

Activities 1 and 2 were designed to yield critical information for improvements and aid in isolating basic material response effects from purely joining based assembly characteristics. This removal of the confounding of several variables was judged essential to both the proper focus of development activity specifics and the establishment of an understanding of the capabilities and limitations of each that could be translated into realistic material and configuration decisions. Activity 3 would serve as a final demonstration to the user office that the material/fabrication technology was fully capable of meeting that user office needs in a manner that did not introduce system degradation or net cost increases. (U)

(U) Initial guidelines, or general goals, were established to serve as both an overall direction for developmental activity focus and to provide a reference point from which it would be possible to evaluate the state of program progress at any given stage of the subsequent phases of the work. These are summarized in Table 1. The most immediate consequence of these criteria was the selection of the Kill Vehicle (KV) structure as representing that portion of the interceptor which would yield the greatest near-term system benefit for the developmental funding available.

(U) It should be apparent that, for a given PROPERTY/PROCESS entry, some items listed as EXAMPLE ANTICIPATED BENEFITS/REQUIREMENT are potentially in contradiction (e.g., a mass reduction may be apportioned in several ways). This represents a deliberate selection by KT of a viewpoint that may be summarized as: "The legitimate role of Key Technologies is to provide alternatives which can be used by the Project Offices to achieve performance gains. Finalized system design trades which lead to the selection of a particular performance gain(s) as the most desirable is solely a function of the Project Office responsible for the given element. Close coordination between Key Technologies and the user Project Office(s) will be maintained to ensure that the developmental activity is matched to the gain(s) selected."

(U) Although it can be argued that an interactive approach somewhat limits the autonomy of the technology directorate, reality has shown that the gains achieved far outweigh any such loss. Access to information critical to increasing the scope and detail of the rather general guidance of Table 1 has been made available. Detailed designs have

been provided when required. Changes in specific performance parameters and configuration modifications have been communicated as they occur, not months later. As a consequence, the quality and timeliness of technology directorate recommendations have been greatly enhanced. In turn, KT generated recommendations for changes which take maximum advantage of the potential of a new material have been given serious review and have been incorporated when these recommendations have proven to be consistent with the overall goals for the element – a condition that has been met in the majority of such cases. Project Office involvement from the earliest stages of the development activity through final qualification testing has had a major program impact: acceptance by the element involved of a portion of the test and evaluation funding load. This enables a more effective use of development funds and created an essentially automatic technology insertion process. (U)

TABLE 1. (U) General guidance for development activities directed toward use of advanced composite materials in USASDC interceptors.

PROPERTY/PROCESS	EXAMPLE ANTICIPATED BENEFITS/ REQUIREMENT
1. MASS REDUCTION	<ul style="list-style-type: none"> • INCREASED SYSTEM MANEUVERABILITY • DECREASED SYSTEM BOOSTER MASS • DECREASED SYSTEM SIZE
2. STIFFNESS/BENDING MODE FREQUENCY/STRENGTH INCREASE	<ul style="list-style-type: none"> • REDUCED BORESIGHT ERROR • REDUCED SENSOR COMPLEXITY/ COST • INCREASED STRENGTH RELATIVE TO PRIMARY LOADING VECTOR
3. CO-LOCATED COMPONENT DEGRADATION (e.g., outgassing induced sensor degradation)	<ul style="list-style-type: none"> • NONE
4. FABRICATION COST	<ul style="list-style-type: none"> • SYSTEM COST CHANGE APPROXIMATELY ZERO

ATHA, ENNIS

(U) The early hardware phase of the effort and those lessons learned from the advanced composite KV structure development for Exoatmospheric Re-entry Vehicle Interceptor System (ERIS) are discussed in the next sections. We follow this with an examination of Ground-Based Interceptor (GBI-X) prototype KV structure development and advances in improved/affordable fabrication technology. This final section summarizes future activities of the Structural Materials Program.

Advanced Composites Application to the ERIS KV Structure (U)

(U) As previously indicated, the initial effort was directed toward development of a composite material substitute for the Aluminum/T300 Graphite /934 Epoxy baseline ERIS KV structure (approximately 1.4 m in length by 0.5 m at the greatest diameter). Metal Matrix Composites (MMC), Thermoset (TS) Resin Matrix Composites (RMC), and Thermoplastic (TP) RMC were examined as candidate categories. A category downselect lead to TS RMC as the primary option and MMC as the secondary choice. The decision was based on several factors: 1) the required materials were readily accessible; 2) both the requisite engineering characterization and experience needed for a high probability of success had been established under previous U. S. Army advanced composites application research; 3) a TP RMC effort was to be initiated under the USAF M&S program, implying leveraging versus duplication as the route which would permit optimum development path for immediate use while retaining the potential for latter application of this material to ground-based interceptor configurations, if warranted. The designation of TS RMC as the primary option was based upon the mass reduction potential of the state-of-the-art (SOTA) version of this material relative to its MMC counterpart. Project Office feedback played a crucial role in the downselection process. TS RMC mass reduction potential was evaluated against a theoretically higher strength and stiffness potential for the MMC alternative. Participation in ERIS project reviews and follow-on ERIS - KT discussions clearly indicated that, in the 1988 - 1989 timeframe, mass constraints presented the greatest challenge to the ERIS system designers. Although not the dominant consideration, the projected cost per unit of a production quantity product also influenced the decision to give the TS RMC option a higher priority. (U)

(U) A view of the primary subcomponents of the ERIS KV structure, along with baseline and alternative materials for each subcomponent, is shown in Figure 1.4. Materials listed after the baseline entry are given in the order of their respective priority for investigation. With the single exception of the Forward (FWD) and aft shells, multiple alternatives are designated for each primary subcomponent. The cruciform design reflects the selection of

the ERIS prime contractor, as approved by the ERIS Project Office. It should be noted that the baseline design material selection was later altered by the ERIS prime contractor to replace those subcomponent materials labeled as graphite/epoxy with 7075-T6 aluminum -- in effect, the development activity and subsequent testing became a direct comparison between a conventional all metal structure and an essentially all composite structure. (U)

(U) The TS RMC version of this design is often referred to as a 1:1 replacement. This term indicates that the external physical dimensions of the structure are such that it could physically replace the baseline with little, or no, adaptation of other KV components required. As discussed below, alterations which are consistent with the mechanical properties of composites and do not affect the overall adaptability of the structure were made when justified by modeling predictions that were supported by test data.

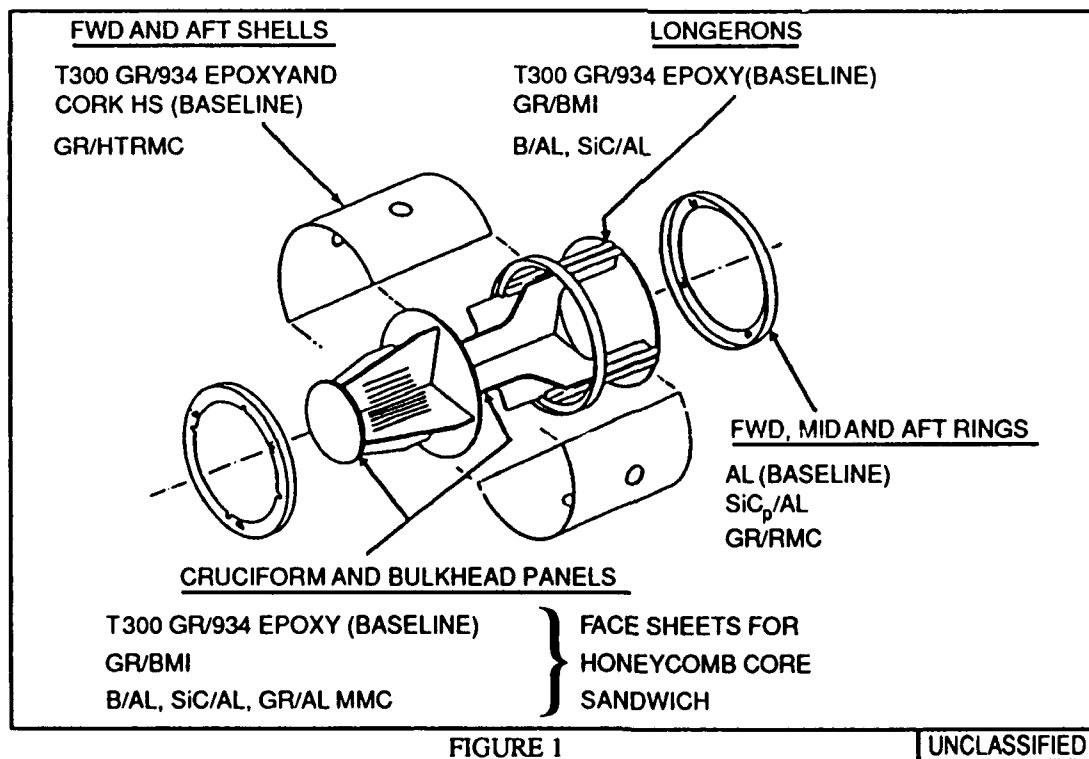


FIGURE 1. (U) Initial ERIS KV assembly: subcomponents, baseline materials, and prioritized composite material evaluation options.

ATHA, ENNIS

(U) A range of potential resins, and associated fibers were examined: bismaleimides (BMI) V398G, F650, and MR56-2; copolymer HLX11 171; epoxy NOVAL AK HRX 159A; polyimide PMR15, IP 600, NR15082; phenolic 908; polybenzimidazole; acetylene terminated polyphenolene (ATP); ATP/polyphen-olquinoxylene. Physical properties used in the trade analysis were selected to focus upon interceptor needs: maximum design temperature; char yield and rate; processability; outgassing; damage tolerance; stability under hot/wet conditions. In addition, the availability and cost considerations previously mentioned were used as both an eliminator factor and to differentiate between options which were otherwise approximately identical in rank. (U)

(U) Non-resin component materials included alumina; silicon carbide; boron carbide. The candidate MMC matrix materials examined were aluminum, magnesium, and titanium – with either continuous or discontinuous reinforcements. The considerations listed in the previous paragraph were applied to selections in these categories.

(U) The RMC trade studies, finite element modeling, and material testing produced a final set of subcomponent materials. A BMI, MR56-2 tape with P100 graphite fiber reinforcement formed in panel configuration over aluminum honeycomb, was selected as the matrix/fiber/core system for both the cruciform and bulkhead elements. The cruciform panels were formed using eight plies of the P100/MR56-2 tape. A zero, plus 45, zero, minus 45 degree layup provided preferential strength and stiffness parallel to the primary axis of the cylindrically symmetric KV structure configuration. The layup thickness was 0.024 inches. Identical materials and ply counts were used for the bulkheads. However, a zero, minus 45, 90, plus 45 degree pattern was chosen to introduce pseudo-isotropic mechanical response. The resin content of the final products was 35 - 40%. The void content was held to less than 1%. Cure/post-cure chemistry was verified via Fourier Transform Infrared (FTIR) spectroscopy.

(U) All secondary structural components were fabricated using a second BMI, V398G with Hitex46-8B graphite fiber reinforcement. The optimum configuration proved to be zero, 45, zero degrees in a 3 ply, 0.042 inch layup. The zero degree filaments provide primary axis stiffness while the 45 degree filaments supply the necessary shear strength to support structure fasteners. Evaluation of resin content yielded values within the range of those given for the cruciform/bulkheads. Void content was also less than 1%. Cure/post-cure chemistry verification again employed FTIR.

(U) As previously stated, a parallel effort focused upon joining and

bonding was included in the overall composite KV structural program. As originally planned, the intersection of the two main panels of the cruciform would be secured with right angled strips. Examined in cross-section, the cruciform divides into four quadrants with one angle piece at every junction, or one per quadrant. One fastener would pass through two angles pieces and one cruciform panel. This yields a total of four fasteners, and any associated reinforcement hardware, per joining station. The attachment and load distribution of the baseline design was not consistent with the behavior of composites under maximum projected loading. The primary difficulty is the introduction of a lowering of the joint strength and long-term assembly stiffness characteristics by the creation of stress which results in direct peel of the joining adhesive. Compensation could be made, but this increased the number of fasteners -- already a major driver of the KV structure mass. (U)

(U) A U-channel design was formulated which resolved the joining difficulty. A cross-section of a fastener station for the U-channel approach shows one panel of the cruciform as continuous. The perpendicular panel is in two parts with a U shaped channel oriented such that the bottom of the U forms the interface with the continuous panel. A single fastener passes through each U-channel and the continuous panel. The net result is that the number of axial support components is reduced from 4 to 2 and the number of fasteners at each station is reduced from 4 to 1. The total number of fasteners required -- relative to the baseline -- is reduced by slightly more than 50%. The gain is even greater when the need for additional fastener stations, associated with the combination of the original baseline design and application of RMC as the KV structure material, is taken into consideration. The mass of two U-channels is 44% lower than that of four angle forms. It was found that potting -- 82-EP-PB Eccolite -- inserted in the continuous panel could replace metal inserts with a mass reduction of 90% per fastener. Titanium fasteners were used to provide high strength, low mass, and galvanic corrosion protection. It should be noted that practical application of this design requires a high preload on the throughbolt and that this preload must be maintained under all flight stresses. With these conditions met, the face sheets are not subject to significant transverse tension loading. Further, the adhesive -- a reticulating FM300U -- is placed in pure shear. A final modification -- the replacement of potting along the edges of the honeycomb substrate by perforated aluminum tape -- further reduced parasitic weight and simplified the problem of ensuring that moisture build-up would not occur.

(U) Preliminary evaluation of the complete TS RMC structure and a joining test article were conducted prior to integration of the shell

ATHA, ENNIS

subcomponents with the full-scale assembly. Non-destructive loading and vibration testing confirmed adequate strength. Modal analysis indicated a 40% increase over the 60 Hz requirement in the lowest bending mode frequency for the completed full-scale assembly. The benefit of primary concern to the ERIS Project Office significantly exceed initial projections: the mass of the metal baseline was 10 kg; preliminary estimates projected a 35% mass reduction for the TS RMC KV structure; the final TS RMC KV structure proved to have a mass 42% lower than the baseline. The role of the bonding and joining technology development in these results cannot be over emphasized. It is well known that the primary obstacle to translating the inherent stiffness and strength of composite materials lies in the inability of joining interfaces to match, particularly under repeated loading, these characteristics. The value obtained represents a major gain in this area. The advance from a 35% mass reduction to a 42% mass reduction is almost entirely due to the sum of the individual bonding and joining mass savings discussed in the previous paragraph. (U)

(U) With the addition of the shell subcomponents, the completed KV structure was ready for modal testing. In accordance with the previous direction from KT and the ERIS Project Office, the assembly was transferred to a common test site for side-by-side evaluation with the ERIS prime contractor developed aluminum baseline KV structure. Dummy loads were attached to simulate the various KV components (e.g., tanks, flight power batteries). Further mass reduction gains were observed: the total TS RMC assembly mass proved to be slightly in excess of 52% lower than the baseline mass. The measured value for the lowest bending mode frequency was 97 Hz – somewhat over 50% higher than the requirement and the baseline structure value – 41% greater than the counterpart component.

(U) An identical MMC structure was also prepared and tested. Boron fiber reinforced aluminum (B/AI) was selected for fabrication of the cruciform panels and bulkheads. A unidirectional, two ply (p.015 inch) layup was used. All secondary subcomponents were fabricated using silicon carbide particulate aluminum (SiC_p/AI). All testing was a direct parallel to that used for the TS RMC and baseline aluminum structure. It should be noted that it was recognized that the MMC version would exceed the mass of the TS RMC KV structure by a significant degree. However, the potential for a major stiffness gain was judged to be of sufficient value to justify the effort. The end results conformed to mass expectations: 12% reduction relative to the baseline value for the bare structure. Although several innovations were made in the machining of the SiC_p/AI, and in MMC joining technology, the overall performance under modal testing did not meet expectations. Post-test analysis indicated that a considerably thinner tape would be required to

achieve both mass and stiffness values that would be competitive with the TS RMC approach. This, combined with the greater projected cost per unit of the MMC KV structure, resulted in a decision to delay further work toward this type of application until such time as the necessary basic research and development had been completed. (U)

Major Lessons Learned (U)

(U) The TS RMC KV structure exceed all initial performance projections. The completed technology represents a significant advance over the aluminum baseline. As a direct result, the Project Office has elected to incorporate advanced composites as an integral part of new baseline KV structure guidance. Cost reduction, a function of the points raised in the remaining paragraphs of this section, remains a potential issue.

(U) The parallel effort in the area of bonding and joining technology development clearly provided major benefits to the main program focus. The first element of this is the obvious increase in the mass reduction gains. Further, the introduction of a well defined, simplified attachment procedure decreased the labor and time required for assembly of the associated components. That this development would yield some benefit was not unexpected. The degree realized did exceed anticipated results. However, by comparison to analogous techniques used with all metal structures, the SOTA for bonding and joining is not such that the full practical potential of the basic composite material is achieved. To reach this goal in a cost effective manner both the mechanical efficiency of TS RMC joints, and the joints for other composites such as TP RMC, require further enhancement. In addition, the number of joints required must be drastically reduced. The latter relates directly to a second area of required improvement.

(U) A parts count summary for the main components of the ERIS KV structure is given in Table 2. In addition to the listed components, fastener assemblies are required. It is clear that, as long as the number of units to be combined to form the final assembly is of this order, the degree of labor required to mate these components and the potential for rejection of an assembly due to joint defects is high. This situation is relatively unimportant in the one-of-a-kind fabrication associated with a prototype. However, cost effectiveness of the end product is a major goal of the entire development program. Thus, alteration of the fabrication approach via reduction of the parts count is clearly required.

(U) The weakening of composites at attachment points due to fiber

cutting/matrix damage is a well known phenomenon. The KV structure must serve as an attachment base for a wide range of KV components. The process of cutting holes, in addition to providing for sealing and reinforcement as required, adds significant labor overhead to the fabrication process. Further, no practical approach was found which would permit complete restoration of the deterioration associated with machining damage -- not an unexpected situation. (U)

TABLE 2. (U) Subcomponent Count Summary for the TS RMC ERIS KV Structure

ITEM	QUANTITY
CRUCIFORM PANELS	
FULL PANEL	1
HALF PANEL	2
BULKHEADS	3
U-CHANNELS	4
U-CLIPS	12
RINGS	3
LONGERONS	8
ANGLES	8
HALF SHELLS	2
TOTAL	43

(U) Each of the above points, combined with the clear-cut performance gains available through the use of advanced composites, lead to the decision to initiate further development. Material selection, processing development which would be consistent with the fabrication changes required to address these points was made the focus of the activity.

Materials and Fabrication Approaches for GBI (U)

(U) USASDC SLKT KT activity to address these issues was initiated in the second half of 1990. The basic concept -- *in situ* sensors to monitor temperature and pressure during the forming process combined with a near-net shape mold to minimize post-cure machining -- is often referred to as "Smart Molding". The near-net shape mold contains pre-positioned forms to provide mounting points. Computer monitoring and control enhance the "Smart" aspect of the process. Rapid progress has been made, with the current end-product meeting or exceeding expectations. However, a

considerable body of work remains. For this reason, the current SOTA might be better characterized as "semi-intelligent" molding with the final goal designated as "Smart". This section examines the initial application of this concept and its extension to generic GBI-X structures.⁵ (U)

(U) The first test structure prepared using the process was a full-scale Brilliant Pebbles (BP) structural bridge. The basic form was that of a simple cylinder with a flange at either end. Approximately 35% of the cylindrical body material was removed to reduce mass. Flat, square mounting surfaces with circular mounting holes were provided near the mid-section of the cylinder. The material combination which proved to yield the best forming/performance match consisted of a hybrid T650-52/BMI and P100/BMI TS RMC set. The article met, or exceeded all stiffness, strength, and bending frequency requirements.

(U) As with the original structural material program, two candidates were selected for the "generic" GBI-X structures. It should be noted that the term "generic" refers to the current lack of a detailed design, not to a vague approximation of the best KT guess as to the size and loading of the final GBI KV structure. The GBI Project Office has provided KT with current contractor concept development data and this has been incorporated into the specifics of the KT effort. It should also be noted that, at this time, there are several competing designs/contractors. As a consequence, some aspects of the structural form and requirements are competition sensitive and will not be reported in this paper. It can be stated that all configurations under development are axially symmetric and are approximately 40% of the size of the ERIS KV structure.

(U) The material selection process was essentially identical to that used for the ERIS KV structure. One additional feature was the adaptability of the candidate material to the fabrication process -- a component of the processability condition previously used, but considerably constrained relative to the ERIS structure. The two categories considered were TS RMC and TP RMC. The final TP RMC selection was T650-42/RADEL 8320. The full-scale structure was formed in five basic steps: compression molding of inner cylinder; diaphragm formation of bulkhead skins; bonding of skins to honeycomb core; bonding of bulkheads to inner cylinder; mechanical drilling of mounting point holes. The predicted lowest bending frequency for this 2.7 kg structure is 270 Hz.

(U) The TS RMC selection was T650-42 and P100 fibers using MR56-2 BMI resin. The full-scale structure fabrication was reduced to two steps: matched metal net molding; tag removal. Matched metal net-shape

ATHA, ENNIS

molding both eliminated all major bonding requirements and eliminated the need to machine mounting point holes. The TS RMC KV structure has a slightly higher mass than that of the TP RMC version -- 3.0 kg. Its predicted natural frequency is 300 Hz. (U)

(U) Both the preceding discussion and the "Lessons Learned" section of this paper would imply that all MMC work has been discontinued. This is not the case. One application -- that of the Tactical, or Theater, Missile Defense (TMD) D-2 Hypervelocity Projectile -- has strength and temperature requirements for its "shell" structure that require the use of an MMC. KT has initiated work similar to that described for BP and GBI, but using B/Al. In this application, the uniform material composition, relatively small size, and simple shape (cone) combine with a copper diaphragm form of matched metal near-net shape molding to offer a reasonably cost effective utilization of MMC materials. The main technical challenges encountered to date involve the extremely sharp taper of the cone.

Future Development Activity (U)

(U) Modal testing of the TP RMC and TS RMC GBI structures is planned for the latter half of 1992. The MMC work will be concentrated upon maintaining structural integrity and the fabrication repeatability under the D-2 design discussed above. As downselections are made, second generation GBI structures will be developed. A major component of this effort is the evaluation of new ultrahigh modulus fibers that are being advanced as having significantly improved properties when deformed (e.g., edge curvature induced bending). Elongation tolerance increases of over 50% have been asserted. If these are to be of utility, the claims must be verified and their adaptability to other conditions of the near-net shape molding must be established. Application of the KT structural work to TMD systems has been set as a major goal. The experience of the last few years has firmly established the mode for execution of these activities: close, responsive interaction with the responsible Project Offices. This approach will be continued and extended.

References (U)

1. *Advanced Structural Materials Technology Development Program -- Materials Characterization Test Plan*, Report Number LJ-86-85-TR (October 1986).
2. *Advanced Structural Materials Technology Development Program -- Joints Test Plan*, Report Number LJ-86-87-TR (October 1986).

ATHA, ENNIS

3. *Advanced Structural Materials Technology Development Program -- Critical Interfaces Test Plan Phase I*, Report Number LJ-86-84-TR (October 1986).
4. *Advanced Structural Materials Technology Development Program -- Interim Technical Report No. 2*, Report Number LJ-89-024-TR (July 1989).
5. *Interceptor Structural Materials -- Progress Report*, Report Number LJ-91-095-TR (December 1991).

Analytic Dispersion Relations Near the Γ -point in Strained Zincblende Semiconductors

Thomas B. Bahder,* Ph.D.

Harry Diamond Laboratories, U.S. Army Adelphi Laboratory Center
2800 Powder Mill Rd., Adelphi, Maryland 20783-1197

1. Introduction

Optoelectronic devices based on thin layers of semiconductor materials are expected to play an important role in the next generation of Army systems. Such devices are typically fabricated by growing thin layers of zincblende semiconductors, such as GaAs and InGaAs, by techniques such as Molecular Beam Epitaxy (MBE). Most semiconductor materials, such as GaAs and InGaAs, do not have the same lattice constant, and consequently, one or both materials have a strain that is built-in during crystal growth. Such built-in strains frequently lead to improved device performance. However, the built-in strain modifies the bulk band structure of the constituent semiconductor materials in a complicated way. This modified band structure serves as the basis for calculating heterostructure electronic states, typically by applying effective mass theory. The details of the electronic states determines the device electrical and optical characteristics. In order to design useful optoelectronic devices we must understand the band structure of the different semiconductor materials under conditions of strain.

We use an eight-band $k.p$ model¹ of the electronic states to accurately describe the band structure in the strained semiconductor near the Γ -point in the Brillouin zone, the region of greatest significance for optical transitions. This model leads to a set of 8×8 matrices, describing the various interaction terms, whose sum must be diagonalized to find the band dispersion relations. The model contains several material parameters which serve to determine the band structure. A practical difficulty with this model is that as we go from one material system to another, at each wave vector k in the Brillouin zone, the matrices must be diagonalized numerically for each set of material parameters. Consequently, it is difficult to see the effect of the material parameters on the band structure and on the resulting device characteristics. In this work we apply Lowdin perturbation theory², together with symbolic manipulation techniques³, to obtain analytic formulas which describe the band structure near the Γ -point. These formulas give the explicit dependence of the energy bands on the material parameters and the strain. The technique we have demonstrated to obtain analytic dispersion

relations is general and it may be applied to other regions of the Brillouin zone, e.g., the X-point, and other classes of materials where explicit band structure formulas are needed.

2. Background

Energy bands in strained semiconductors, such as Si and Ge, have been the subject of several theoretical investigations.⁴⁻⁸ More recently, the direct-gap semiconductors with the zincblende structure, such as GaAs, InAs, and GaP, are attracting considerable attention because of their demonstrated applications in the area of optoelectronic devices.⁹⁻¹¹ Energy bands in strained semiconductors are often calculated by $\mathbf{k}\cdot\mathbf{p}$ type models, leading to matrices whose dimensionality depends on the small point group of \mathbf{k} (the appropriate symmetry group at the critical point in the Brillouin zone), and the number of bands that are included in the model.¹² For example, near the Γ -point in the Brillouin zone of a direct-gap zincblende semiconductor, the band structure in the vicinity of the fundamental gap is accurately described by an eight-band $\mathbf{k}\cdot\mathbf{p}$ model, whose basis states form basis functions for the T_d double group.^{12,14-16} This formulation includes effects resulting from the lack of inversion symmetry but leads to 8×8 matrices which must be diagonalized numerically. In this work we formulate a renormalization procedure based on Lowdin perturbation theory² which allows one to derive analytic dispersion relations in terms of the parameters in the original $\mathbf{k}\cdot\mathbf{p}$ model. As a concrete example we demonstrate this renormalization procedure using an eight-band $\mathbf{k}\cdot\mathbf{p}$ model¹ of the energy bands near the Γ -point of strained zincblende crystals. The model includes the conduction band, the (light- and heavy-hole) valence bands and spin-orbit split-off bands. Lowdin perturbation theory is used to derive the 8×8 renormalized interaction matrix, $U = H + D$, which describes the band structure near the Γ -point. The matrix H is essentially the eight-band Kane model^{1,13} and is independent of strain, and D contains the strain interactions to first order.¹ We use Lowdin perturbation theory to perform a second renormalization on the matrix U , leading to matrices of smaller dimensions, 2×2 , 4×4 and 2×2 , for the conduction, light- and heavy-hole, and spin-orbit bands, respectively. These second-renormalized matrices describe the band structure near the Γ -point and are functions of the material parameters that appear in the original eight-band model. The reduced dimensionality of these matrices allows us to obtain analytic results for dispersion relations near $\mathbf{k}=0$, in the presence of arbitrary strain. In what follows, we outline how these second-renormalized matrices and resulting dispersion relations are calculated.

3. Second Renormalization

In the strained eight-band model the conduction band states at $\mathbf{k}=0$ form a basis for the Γ_6 irreducible representation of the T_d double group.^{1,15} The states at $\mathbf{k}=0$ for the light- and heavy-holes, and for the spin-orbit split-off states, form a basis for the Γ_8 and Γ_7 irreducible representations, respectively. See Ref. 1 for definitions of these basis states. In this basis, the total Hamiltonian matrix of the eight-band model is given by $U = H + D$, where H is essentially the eight-band Kane model (independent of strain) and D is the sum of the strain interaction matrices, whose elements are linear in strain. The 8×8 matrices, H and D , are given in Ref. 1. The matrix H depends on the modified Luttinger parameters, γ_1 ,

γ_2 , and γ_3 (which are functions of the Luttinger parameters,¹⁶ γ_1^L , γ_2^L , and γ_3^L , the valence-

conduction band mixing parameter, P_o , the mixing with higher bands outside of the Γ_6 , Γ_8 , and Γ_7 manifold of states, A' , and the energy gaps at $k=0$ (between the conduction and valence band $E_g = E_c - E_v$, and between the valence and spin-orbit band, Δ). The strain interaction matrix, D , depends linearly on strain tensor components, e_{ij} , and on the deformation potential constants a' , b' , a , b , and d . The deformation potential constants a' and b' describe the coupling of the conduction band to strain and a , b , and d , are the Pikus and Bir^{4,7} deformation potential constants describing the coupling of the valence band to strain. See Ref. 1 for precise definitions of these deformation potential constants in this work.

We now perform a second Lowdin renormalization on the matrix U . For the conduction band, we treat the Γ_6 conduction band basis states as class A states and the Γ_8 and Γ_7 basis states as class B states.² The second-renormalized conduction band matrix, $U_{cc'}$, can then be written as a sum of two 2x2 matrices $U_{cc'}^{cond} = U_{cc'}^{(o)} + U_{cc'}^{(1)}$ where $U_{cc'}^{(o)}$ describes the conduction band in the absence of strain and $U_{cc'}^{(1)}$ describes the renormalized strain interactions, with matrix elements linear in strain components. Third-order Lowdin renormalization leads to

$$U_{cc'}^{(o)} = H_{cc'} + \sum_n^B \frac{H_{cn}H_{nc'}}{(\bar{E} - H_{nn})} + \sum_n^B \sum_{n'}^B \frac{H_{cn}H_{nn'}H_{n'c'}}{(\bar{E} - H_{nn})(\bar{E} - H_{n'n'})} \quad (1)$$

where the elements of $U_{cc'}^{(o)}$ are correct to fourth order in wave vector.¹⁷ The indices "c" and "n" refer to class A and class B states, respectively. Matrix elements $H_{cc'}$ and $H_{nn'}$ are second order in wave vector, while matrix elements that connect class A and class B states, H_{nc} are linear in wave vector. The second-renormalized strain interaction matrix is given by

$$U_{cc'}^{(1)} = D_{cc'} + \sum_n^B \left(\frac{H_{cn}H_{nc'}}{(\bar{E} - H_{nn})} \frac{D_{nn}}{(\bar{E} - H_{nn})} + \frac{H_{cn}D_{nc'} + D_{cn}H_{nc'}}{\bar{E} - H_{nn}} \right) \quad (2)$$

where the elements of $U_{cc'}^{(1)}$ are correct to linear order in strain and quadratic order in wave

vector. The first term in Eq. (2) represents the direct strain interactions among the Γ_6 states and the second term represents the renormalization of the conduction band effective mass due to the presence of strain. The above renormalization procedure is reasonable for wide gap semiconductors such as GaAs since, for small wave vectors and small strain, the off-diagonal elements are small compared to the band gaps at¹⁸ $\mathbf{k}=0$, $|H_{cn}| \ll |H_{cc} - H_{nn}|$. We linearize the

eigenvalue problem for the conduction band by taking $\bar{E} = (H_{11} + H_{22})/2$. Since the wave vector, \mathbf{k} , and strain, e_{ij} , are both small quantities, the matrix elements of $U_{cc'}^{(cond)}$ are

expanded in a power series in wave vector and strain components. The conduction band dispersion relations are then found from the eigenvalues of the resulting 2x2 matrix.

For the light- and heavy-hole valence bands the second-renormalized matrix is defined the same way as in Eqs.(1)--(3) except that the Γ_8 states are taken as class A states and the Γ_7 and Γ_6 states are taken as class B states. In this latter case we take $\bar{E} = (H_{33} + H_{44} + H_{55} + H_{66})/4$. The spin-orbit band is treated in an analogous fashion. In

this way explicit analytic dispersion relations for the conduction, light-hole, heavy-hole and spin-orbit bands were obtained, in the presence of an arbitrary strain. The expressions for these dispersion relations are long and we do not write them out here. The full details are given elsewhere.¹⁸

In order to carry out this renormalization procedure a substantial amount of analytic computation was required. The symbolic manipulation program, *Mathematica*, was used for the computation.³ The calculations were done on a MIPS RC3240 computer with 40MB of RAM and took approximately 3 hours of cpu time.

4. Conclusion:

Using material parameter values that are characteristic for GaAs and typical values of heterostructure built-in strain ($\sim 10^{-3}$), comparison of the analytic dispersions to those obtained by numerically diagonalizing the eight-band model shows that the analytic dispersions give a reasonable approximation near $\mathbf{k}=0$.

Our results are appropriate for the zincblende group of semiconductors which have a direct energy gap at $\mathbf{k}=0$, for which the small point group of \mathbf{k} , at $\mathbf{k}=0$, is T_d (which lacks inversion symmetry). However, the results for the light- and heavy-hole and spin-orbit bands apply equally well to the case of Si and Ge, which have O_h symmetry (and have inversion symmetry), if we take to zero the terms that must vanish when inversion is present, $B=b'=0$, and $P_0=0$. In the eight-band model we simply neglect the conduction band block. This leads to a six band model without a conduction band appropriate for Si and Ge.

The renormalization scheme we have used is quite general and consequently can be applied to other many-band $\mathbf{k}\cdot\mathbf{p}$ models to derive analytic dispersion relations.

Acknowledgements

I thank John D. Bruno for many helpful discussions.

References

1. T. B. Bahder, Phys. Rev. B **41**, 11992 (1990).
2. P. Lowdin, J. Chem. Phys. **19**, 1396(1951).
3. S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer* (Addison-Wesley, New York, 1991).
4. G. E. Pikus and G. L. Bir, Sov. Phys. Solid State **1**, 1502 (1960).
5. See also: G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (J. Wiley and Sons, New York, (1974).
6. H. Hasegawa, Phys. Rev. **129**, 1029 (1963).
7. G. L. Bir and G. E. Pikus, Sov. Phys. Solid State **3**, 2221 (1962).
8. K. Suzuki and J. C. Hensel, Phys. Rev. B **9**, 4184 (1974)
9. H.-R. Trebin, U. Rossler, and Ranvaud, Phys. Rev. B **20**, 686 (1979).
10. H.-R. Trebin, B. Wolfstadter, H. Pascher, and H. Hafele, Phys. Rev. B **37**, 10249 (1988).
11. P. Pfeffer and W. Zawadzki, Phys. Rev. B **41**, 1561 (1990).
12. See for example, F. Bassani and G. P. Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon Press, New York, 1975).
13. E. O. Kane, in *Semiconductors and Semimetals*, ed. by R. K. Willardson and A. C. Beer, vol. **1** (Academic Press, 1966), and E. O. Kane, in *Handbook on Semiconductors* ed. by T. S. Moss, (North-Holland 1982).
14. M. H. Weiler, in *Semiconductors and Semimetals*, ed. by R. K. Willardson and A. C. Beer, vol. **16**, (Academic Press, 1966).
15. We use the notation in G. F. Koster, J. O. Dimmock, R. G. Wheeler, and H. Statz, *Properties of the Thirty-Two Point Groups*, M.I.T. Press, Cambridge, 1963.
16. J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).
17. The original strain-independent eight-band model, H, is correct only to terms $O(k^2)$. However, fourth-order terms that couple the conduction band states via intermediate valence

BAHDER

band states are zero, since $\langle v_i | \mathbf{k} \cdot \mathbf{p} | v_j \rangle = 0$, where v_i are the Class A valence band states at $\mathbf{k}=0$ in the original eight-band model. There are non-zero fourth-order terms which couple the conduction band states via Class B states (in the original eight-band model), but these are small because of the large energy denominators between the conduction band states at $\mathbf{k}=0$ and the class B states.

18. T. B. Bahder, Phys. Rev. B **45**, 1629 (1992).

BAKER

Lightweight Exo-Atmospheric Projectile
(LEAP)

Paul K. Baker, Mr.
U.S. Army Strategic Defense Command
P.O. Box 1500
Huntsville, AL 35807-3801

I. Introduction

The Gulf War demonstrated for the first time the ability of an interceptor system to successfully intercept incoming ballistic threats. As the threat becomes more sophisticated, the capabilities of the interceptors must grow without a significant increase in cost. One of the drivers in the costs of the system is the size and weight of the interceptor and its components. With smaller, lighter systems, the logistics and support needed will be greatly reduced while transportability and mobility will be increased. The goal of the Strategic Defense Initiative Organization (SDIO) Lightweight Exo-Atmospheric Projectile

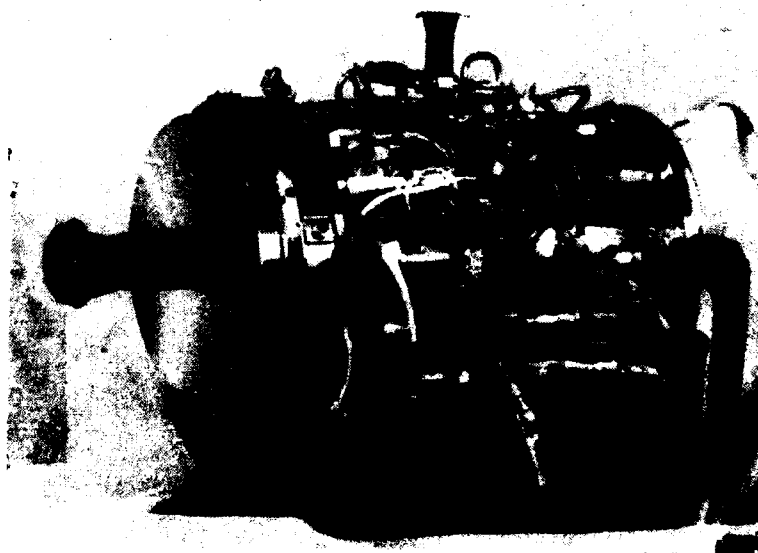


Figure 1 - Lightweight Exo-Atmospheric Projectile (LEAP)

BAKER

(LEAP) program has been the development and demonstration of advanced technologies for light weight, kinetic energy interceptors that will fit the needs described above. LEAP has been able to significantly reduce the weight of interceptor component technologies and integrate them into very light, high performance interceptor technology demonstrators. The Army version, shown in Figure 1, is the smallest and lightest of the LEAP vehicles being developed. The interceptor is only 15 cm (6 in) in diameter by 36 cm (14 in) in length and weighs 6 kg (13 lb) in the test configuration. The program has recently completed a major milestone by demonstrating operation of a fully integrated vehicle in a ground test environment. This testing, including a live strapdown test and a free flight hover test, was accomplished with 100% success. All subsystems performed much better than required. These tests demonstrated the ability of the LEAP body fixed seeker and guidance systems to control vehicle operation in the presence of propulsion system thruster activity with no degradation in overall performance.

The Army LEAP program is managed for the SDIO by the U.S. Army Strategic Defense Command in Huntsville, AL and is executed and technically administered by the U.S. Army Armaments Research, Development and Engineering Center, Dover, NJ. The prime contractor is Hughes Aircraft Company Missile Systems Group in Canoga Park, CA.

II. Overview

The Army LEAP interceptor is an integrated combination of the latest in seeker, propulsion and electronics technologies. The guidance unit, located at the front of the interceptor, consists of the strapdown seeker and flight electronics. The imaging infrared seeker has beryllium primary and secondary reflective optics with a 128 x 128 pixel Mercury Cadmium Telluride (HgCdTe) focal plane array (FPA) housed in a light weight dewar. The electronics unit, located directly behind the primary mirror, is a 14 cm (5.5 in) electronics card formed by mounting eight wafer scale daughterboards on both sides of a single motherboard. The daughterboards each perform a specific function and help off load the central processor. This architecture allows the use of a standard Intel 80386 processor with 4.2 MIPS throughput giving the interceptor added flexibility by being programmable for a variety of targets. The integrated guidance unit weighs in at only 864 gr (1.9 lb).

The propulsion system, developed by The Marquardt Company, is a liquid hypergolic system using hydrazine fuel and nitrogen tetroxide oxidizer. Four 160 Newton (36 lb) divert engines located at the center of gravity (CG) of the interceptor provide lateral maneuver capability to intercept the target. Attitude control is provided by eight 4.5 Newton (1 lb) warm gas thrusters located on the aft bulkhead of the interceptor. The warm gas for attitude control thrust is developed by diverting some of the fuel through a catalyst bed in a warm gas generator located on the front bulkhead. Three small gas bottles on the front bulkhead hold 14,000 psi of

BAKER

helium to pressurize the propulsion system. Total divert capability of the interceptor is greater than 500 meters/second (1640 ft/sec) in the test configuration.

Located on the aft bulkhead are the other elements of the LEAP; including the thermal battery, inertial measurement unit, and the 11.2 Mbit/second telemetry system.

III. Ground Test Program

The ground test program was developed to validate the advanced technologies being integrated into the interceptor before the space flight testing was attempted. The program was structured to demonstrate each technology at the component, subsystem, and finally, the integrated interceptor level. The subsystem and system testing were completed with all performing better than predicted. The next step was the integration of complete interceptors for full up testing. Three interceptors have been integrated to date. The first interceptor, configured for ground test operations, successfully demonstrated the performance of the LEAP concept in a fully operational strapdown test and a free flight hover test at the SDIO National Hover Test Facility. Descriptions of the tests and the results are summarized below.

Two more interceptors in a Space Test Projectile (STP) configuration have recently been completed. These projectiles will be used to further validate the LEAP advanced technologies by demonstrating the interceptor performance in a space environment against realistic targets. Three additional interceptors are currently in fabrication for a total of six flight units including one ground test projectile and five space test projectiles.

Strapdown Test

Both the strapdown and free flight hover tests were conducted at the National Hover Test Facility at Edwards Air Force Base, CA. The strapdown test was the first integrated system firing of a fully integrated and functional LEAP and served as a dry run for the hover test count down and procedures. For the first time, all systems in the interceptor operated while exposed to the thermal, vibrational, and electrical interactions from the other systems. The primary area of interest during this test was the interaction between the guidance unit and the propulsion system. Priority was placed on the instrumentation and the data collection at that interface. Figure 2 shows the instrumented projectile mounted in the static test support fixture during the strapdown system test.

The strapdown test was a safety prerequisite for the hover flight but also served as a major check of the procedures and count down to be used during the hover and upcoming flight tests. In the 10 second test, the

BAKER

LEAP fired a programmed series of divert and attitude control pulses predetermined from simulations. The interceptor was entirely self contained; no commands were uplinked from the ground during the test. The LEAP was shock mounted in a fixture which allowed free structural dynamics to occur but allowed no significant lateral motion and limited rotational motion. The interceptor was fully active during the test and tracked a target located 110 meters downrange (the same one used in the subsequent hover test) and calculated navigational parameters on board, such as position and attitude. A complete telemetry data stream, which included full frame IR sensor video on every seeker data frame, was transmitted from the interceptor to a ground receiving station. External instrumentation was added to measure critical temperatures, pressures, currents, voltages, and structural accelerations. The strapdown test provided more pure engineering data than the hover test due to the addition of this instrumentation. This data proved invaluable for assessment of the integrated system performance prior to committing the projectile to the free flight hover test.

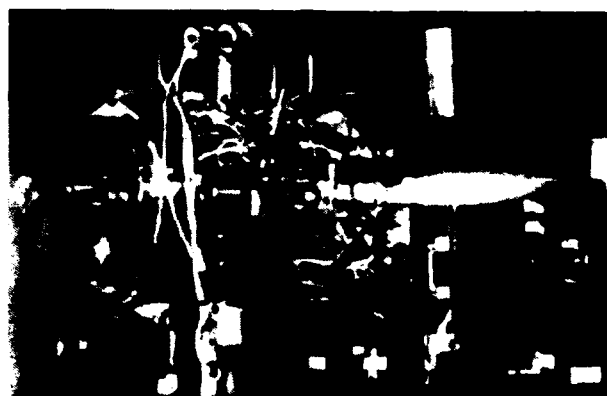


Figure 2 - Instrumented LEAP during strapdown test

Propulsion thrust levels of the integrated system were a critical parameter to be measured during the strapdown test. Due to the small size of the divert and ACS engines it was impossible to directly measure the chamber pressure for thrust level determination. Therefore, the top divert thruster and one attitude control engine were removed during the test and replaced with pressure transducers. This allowed real time measurement of the pressures in the fuel, oxidizer and attitude control system manifolds. The thrust levels were calculated from these pressures based on historical test data from previous thruster component firings.

The downward divert engine was fired on a 5 Hz duty cycle to simulate the required firings to resist gravity during the hover test. With additional firings of the two lateral thrusters, the divert engines were fired a total of 100 times during the test with no pulse to pulse

BAKER

degradation. The attitude control engines were fired 1400 times with no significant warm gas supply pressure variations. The seeker acquired the target and maintained lock throughout the test. Table 1 summarizes the major system parameters that were monitored and compares the results to the requirement.

The interceptor performance, measured against the test requirements below, indicates the strapdown test was a complete success. All test parameters critical to the success of the interceptor in future applications met or exceeded requirement limits. The test data obtained from the strapdown was used to update the LEAP simulation and to calculate expected hover test trajectory and performance. As a result, confidence that the hover test would be successful was very high.

TOPIC	REQUIRED	MEASURED
Passive CG Control	< 2.0 mm	< 0.2 mm
Divert Thrust Levels	143 ± 14 N	143 ± 14 N
Pressure Control Accuracy	1200 - 1280 psi	1200 - 1280 psi
Electronics Unit Temp	≤ 90 °C	84 °C
Tracker Performance	≥ 95 % Valid Track Gates	100 % Valid Track Gates
Structural Dynamics	< 3 GRMS	2.8 GRMS
Structural Bending Frequency	1500 Hz	1500 Hz

Table 1 - Strapdown Test Data Summary

Free Flight Hover Test

The hover test was the next logical step in the validation progression and was the first free flight of the Army LEAP interceptor. During the hover test, the same test setup was used as during the strapdown. This time, however, the LEAP was free to lift off the platform and hover in free flight. The hover test implemented a lock on before launch of the target used during the strapdown. In the 7 second test, the LEAP acquired the target, fired its bottom divert engine to gain an altitude of 3 meters above the launch cradle, and then maintained that altitude for the remainder of the test. At the end of the test, the LEAP shut down the propulsion system, fell into a safety net, and vented the

BAKER

helium pressurization gas to save the vehicle for recovery. Throughout the test, the LEAP tracked the remotely located target while firing its divert engines to maintain an "intercept" course. Telemetry data was relayed to the ground receiving station through the slender fiber optic cable seen trailing the projectile in Figure 3 below. The protective cage was added to the interceptor structure to protect the sensitive optics and propulsion system components from damage at the end of the test. This allows reusing the interceptor for future hover tests.

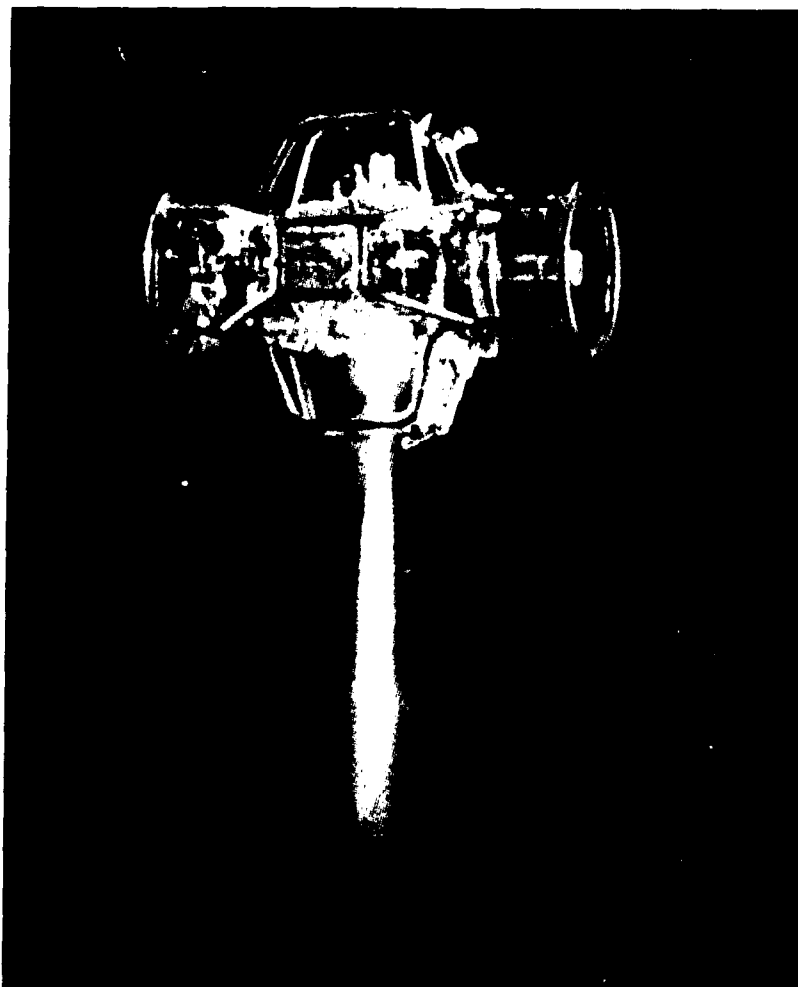


Figure 3 - LEAP in free flight during hover test

BAKER

The fundamental objective of the hover test was to demonstrate the overall controllability of the LEAP while subjected to the harsh propulsion system environment in free flight. The method chosen for implementing attitude control on the interceptor was a major factor in the very light weight system that was achieved. The attitude control engines are very small with a very fast response time (< 1 msec) to maintain stable target pointing within the seeker field of view during flight. While this allows for a very light weight system, the small thrust level and short moment arm do not allow much margin for center of gravity (CG) shift as the propellants are consumed during flight. The divert engines are very large compared to the attitude control engines and therefore could have a devastating effect on interceptor control if they impose too high a disturbance torque on the system when fired. This impact is minimized by locating the center of gravity of the interceptor in the divert engine plane and maintaining it there.

Two methods have been developed to maintain the location of the CG within the specified ± 2 mm requirement. The primary method involving passive CG motion control was employed during the hover test. The fuel and oxidizer tanks in the projectile are paired and located 180° apart. Each pair is drained from opposite ends by driving movable pistons inside the tanks in opposing directions to expel the liquid propellant. The CG is maintained at the center point as long as the expulsion rates from the opposing tank pairs are equal. During propulsion integration testing, tank expulsion rates were tuned and the resulting matched pairs used for the fuel and oxidizer tanks. This method was verified during the strapdown test by measuring the CG of the interceptor both before and after the test. The CG location during that test only changed by 0.2 mm from the pretest measurement; only 10% of the allowed 2 mm range.

The second method developed for CG control is an active CG estimator and controller. In this method, CG location is estimated by the interceptor by monitoring the inertial measurement unit (IMU) outputs during divert thruster firings. If the CG excursion exceeds a set boundary, CG control valves on the propellant tanks are opened or closed accordingly to drain the proper tank(s) and move the CG back to the divert plane. During the hover test, the estimator software was operating for verification, but the controller was not implemented.

The on-board center of mass (CM) estimator calculated that the CG did not migrate more than 0.4 mm in any direction during flight, and was less than 0.3 mm from the divert plane for most of the hover flight. These estimates were verified by CG location measurements taken before and after the hover test. The minimal CG movement is also reflected in the performance of the attitude control system during the test. The target migrated an average of only ± 4 pixels in the seeker field of view during the test and body rates were held to ± 3 degrees per second. This is significantly lower than the ± 20 pixels and ± 8.6 degrees per second requirements.

BAKER

The data collected during flight and subsequently evaluated confirms the LEAP hover flight performance as a complete success. The LEAP performed as expected with no anomalies. Table 2 lists the major areas validated during the test.

The data obtained from the hover test has now been used to further update the interceptor simulation allowing more accurate performance predictions for the upcoming space flight tests. The results show that performance required during space testing is well within the demonstrated performance of the Army LEAP.

TOPIC	REQUIRED	MEASURED
Attitude Control	$\pm 0.17^\circ$	$\pm 0.03^\circ$
Body Angular Rates	± 8.6 deg/sec	± 3 deg/sec
Terminal Guidance (Miss Distance)	± 2 meters	± 0.2 meters
Passive CG Control	$<\pm 2.0$ mm	0.2 m
CG Est Accuracy	± 0.5 mm	± 0.5 mm
Divert Thrust Levels	143 ± 14 N	143 ± 14 N
Pressure Control Accuracy	1200 - 1280 psi	1200 - 1280 psi
Electronics Unit Temp	$\leq 100^\circ\text{C}$	81°C
Tracker Performance	$\geq 95\%$ Valid Track Gates	100% Valid Track Gates
Seeker Noise	≤ 10 Counts RMS	≤ 6 Counts RMS

Table 2 -Hover Test Data Summary

IV. Flight Test Program

The flight test program is a continuation of the validation of the LEAP technologies. Four flight tests are currently being planned with the Army LEAP interceptor. During the flight test program, closing velocities will be increased from test to test to attain progressively more realistic

BAKER

closing scenarios against cold body targets. In this way, the performance of the LEAP technologies can be demonstrated in a near operational environment. The flight test program will begin with two flights at White Sands Missile Range (WSMR). Two later flights, whose increased closing velocities and the complex trajectories cannot be safely flown at WSMR, will be conducted at the U.S. Army Kwajalein Atoll (USAKA) facility.

In the first WSMR flight, the entire experiment package, consisting of the LEAP interceptor, a target, and a support module called the Payload Module Bus (PMB), is carried aloft by a single booster. The entire experiment is performed exo-atmospheric after separation from the burned out booster. The target is released from the PMB with separation boost provided by a cold gas thruster. When the thruster is exhausted, the target reorients itself to point back toward the PMB/LEAP. Several hundred seconds pass while the target/PMB separation grows to approximately 20 kilometers. The target then fires a solid axial thruster to boost itself back towards an intercept point 600 meters to the side of the PMB trajectory with a relative closing velocity of 840 m/sec (nominal). At approximately 17 kilometers separation, the LEAP begins acquisition of the target. It is ejected three seconds later, engages its lateral divert thrusters, and maneuvers toward a hit to kill intercept. During this and all subsequent flight tests, all avionics and full seeker video is being transmitted from the interceptor to the ground over an 11.2 Mbit/sec telemetry data link. This allows complete post test reconstruction of the flight and high fidelity evaluation of actual measured flight performance.

In the second WSMR flight, separate boosters will be used to carry the target and experiment module aloft. This will increase the closing velocity at intercept to approximately 2.5 kilometers/second. Acquisition range will be increased to 34 kilometers to allow time for intercept maneuvers after LEAP ejection.

The final two flights currently planned are being performed at USAKA and utilize two boosters as does the latter WSMR flight. These advanced flights, as depicted in figure 4 below, will have further increased relative closing velocities of approximately 5.5 km/sec for the third flight test, with 7 km/sec closing velocity envisioned for the fourth flight.

In order to meet the highly stressing nature of the advanced space flight experiments, design enhancements are being incorporated into the current LEAP design. A long wave infrared seeker system will be substituted for the current medium wave system on both USAKA flights. This improvement will allow the LEAP to achieve much greater and more realistic acquisition ranges than on the earlier flights with the current design.

The high closing velocities desired by the SDIO for the fourth test of the series calls for the addition of an axial boost stage on the LEAP. The miniature axial kick stage, supplied by the SDIO, will be controlled by the LEAP interceptor during its flyout. This integrated flight system will

BAKER

help to validate the midcourse interceptor concept. To support the additional flight accuracy needed to accomplish this midcourse mission, the current IMU is being replaced with a fiber optic gyro based IMU from Smiths Industries. In conjunction with the IMU upgrade, the electronics unit is being enhanced with a second data processor to complete the midcourse guidance navigation and control the axial booster. As part of the integrated midcourse interceptor system, an RF uplink is being added to receive target state vector updates during midcourse flyout. Guidance algorithms are being added for the midcourse flight and to perform target acquisition after separation of the LEAP interceptor. All previous tests relied on accurate PMB pointing to perform target acquisition prior to STP ejection.

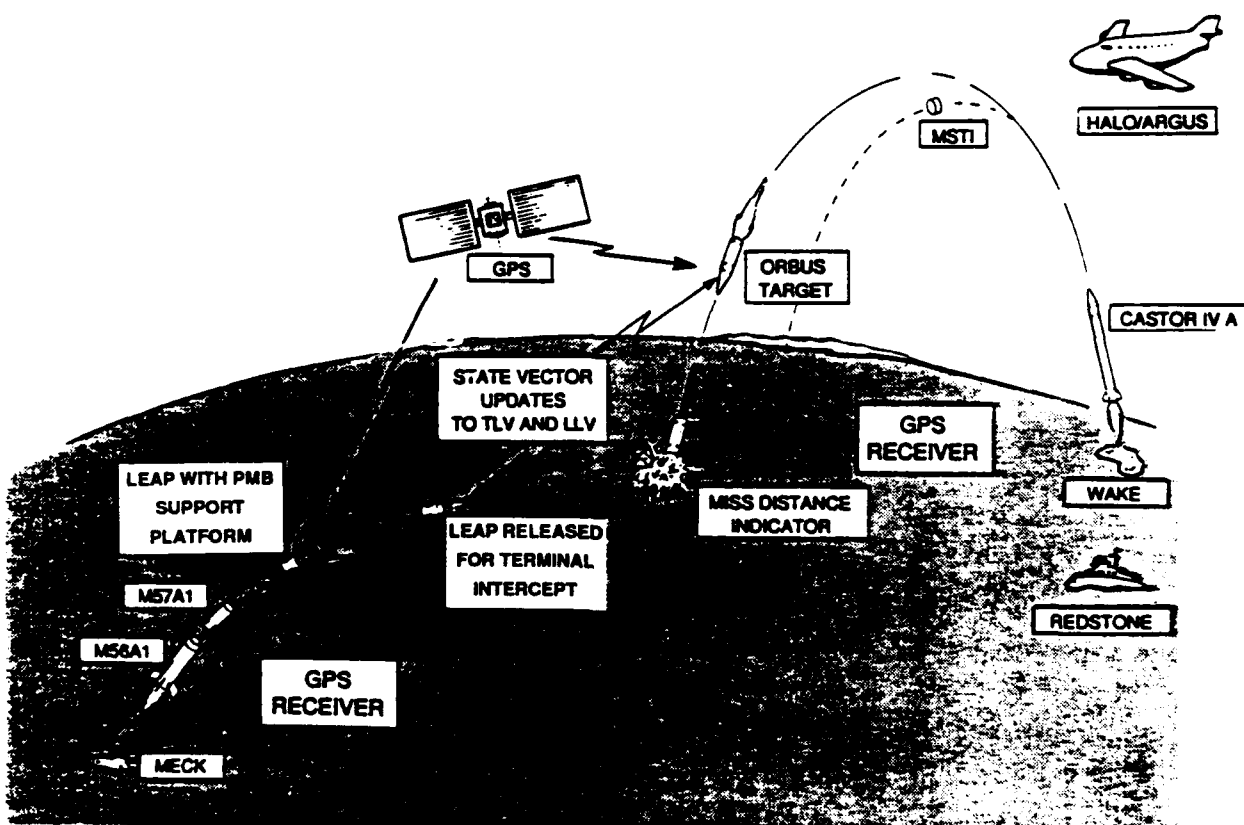


Figure 4 - LEAP advanced space flight mission overview

BAKER

V. Summary

The LEAP program is well on its way to validating advanced technologies for light weight interceptor applications. The successful completion of the strapdown and free flight hover tests demonstrate that the LEAP is ready to enter the flight test phase of the program. The LEAP performance during those tests exceeded all the requirements necessary for successful completion of the flight test series. The validation of the design enhancements being incorporated into the LEAP for the advanced flights will maintain LEAP on the leading edge of technology. These advanced flight test missions will demonstrate the potent lethality of available light weight interceptor technology for hit to kill defense against missile attacks.

BARROWS, FORCH, BEYER, COHEN, MIZIOLEK, NEWBERRY

Development of Laser-Based Ignition Systems for Large Caliber Guns

Dr. Austin W. Barrows*, Dr. Brad E. Forch, Dr. Richard A. Beyer
Mr. Arthur Cohen, Dr. Andrzej W. Miziolek and Ms. Joyce E. Newberry
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

INTRODUCTION

Lasers play an extremely important role in a wide variety of Department of Defense (DOD) applications which include range finders, communications, electro-optics, guidance systems for rockets and missiles and as a battlefield directed energy source. The unique characteristics of lasers also make them excellent candidates as igniters for Army large caliber gun systems. Although the concept laser ignition is not new,¹ recent laser ignition work at the Ballistic Research Laboratory (BRL) has served as the motivation to explore new aspects of this technology. The Laser Ignition in Guns, Howitzers and Tanks (LIGHT) Program was established as a result of a novel ignition concept which was discovered at BRL called **Resonance Laser Microplasma Ignition**.²⁻⁸ It was found that reactive gaseous mixtures could be ignited into combustion very efficiently, i.e., with laser energies less than 1 mJ. A new laser ignition mechanism was formulated based upon the efficient and well controlled resonant formation of laser-produced microplasmas. When the laser is tuned to resonance absorption transitions it was found that as much as sixty times less laser energy was required for ignition. The attractive feature of this ignition source clearly lies in the efficiency of the process which may allow for the development of small low-energy lasers to be used as igniters for energetic solid materials such as gun propellants. A laser source which is tuned to absorption transitions in solid materials or in pyrolysis gases produced at the solid-gas interface could lead to efficient, low energy ignition thresholds. The ability to directly ignite propellant beds could lead to the elimination of primers and igniters from the ignition train which would dramatically minimize vulnerability, simplify the ignition train and facilitate the ignition of insensitive munitions, such as low vulnerability propellant ammunition (LOVA), which are inherently difficult to ignite.

The overall goal within the LIGHT Program is to incorporate lasers into the ignition train. Within the LIGHT Program, ignition has been categorized into two regimes called direct and indirect laser-based ignition. The direct laser ignition concept (the long-term program goal) focuses on initiation of propellant beds via the interaction of laser light with the charge. Indirect laser ignition (the short-term program goal) involves the removal of current primers and igniter material from the ignition train in their present configurations within the munition. Here, the laser light is first transmitted to a sensitizer which then transfers the ignition stimulus to the propellant bed. Both laser ignition concepts involve the transfer of laser radiation into the gun through the use of optical fibers connected through a breech window made of sapphire which will readily transmit the laser radiation with no optical damage and, in addition, withstand the high pressures and temperatures encountered in large caliber guns. Problems associated with window contamination have also been solved.⁹ The breech window can be shielded from the gun combustion using a debris trap and/or cleaned with a breech brush. The laser (attached to the breech or gun mount) must be sufficiently sturdy to survive the high energy gun recoil forces. There are many important characteristics of the laser which must be addressed. These laser parameters include energy, pulse length, wavelength and ignition delay. There are many types of lasers which may serve as candidate igniters and a particularly attractive source is the solid-state Nd:glass laser which operates near 1.05μ , which can be made very small, rugged, reliable, long-lived and inexpensive. Laser radiation near 1μ can readily be transmitted through durable and inexpensive fused silica optical fibers over great distances with negligible loss. This laser wavelength is also readily transmitted through sapphire breech window material. The Nd:glass laser has been extensively used as an ignition source within the LIGHT Program as a result of these attributes.

The aforementioned discussion summarizes key issues which have been considered in the ongoing development of a laser-based ignition system for large caliber guns. The bulk of this paper describes the experimental research and testing in the development of a laser-based ignition system for the **Advanced Field Artillery System (AFAS)** 155 mm howitzer (Unicharge and Liquid Propellant (LP)) and the **Advanced Tank Cannon System (ATACS)** 140 mm tank gun round. The goal of the Unicharge laser ignition program at present is to eliminate the current M82 primer from the gun breech and to ignite the blackpowder igniter pad with a laser with no modification to Unicharge. The focus of the LP laser ignition program is to replace or augment the electrical staged ignition system. The long-term future goal is to eliminate all igniter material from the ignition train and to directly ignite the propelling charge with a laser at multiple points to achieve isochronic ignition. The immediate goal of the ATACS laser ignition program is to replace the primer with an optical feed

BARROWS, FORCH, BEYER, COHEN, MIZIOLEK, NEWBERRY

The overall goal within the LIGHT Program is to incorporate lasers into the ignition train. Within the LIGHT Program, ignition has been categorized into two regimes called direct and indirect laser-based ignition. The direct laser ignition concept (the long-term program goal) focuses on initiation of propellant beds via the interaction of laser light with the charge. Indirect laser ignition (the short-term program goal) involves the removal of current primers and igniter material from the ignition train in their present configurations within the munition. Here, the laser light is first transmitted to a sensitizer which then transfers the ignition stimulus to the propellant bed. Both laser ignition concepts involve the transfer of laser radiation into the gun through the use of optical fibers connected through a breech window made of sapphire which will readily transmit the laser radiation with no optical damage and, in addition, withstand the high pressures and temperatures encountered in large caliber guns. Problems associated with window contamination have also been solved.⁹ The breech window can be shielded from the gun combustion using a debris trap and/or cleaned with a breech brush. The laser (attached to the breech or gun mount) must be sufficiently sturdy to survive the high energy gun recoil forces. There are many important characteristics of the laser which must be addressed. These laser parameters include energy, pulse length, wavelength and ignition delay. There are many types of lasers which may serve as candidate igniters and a particularly attractive source is the solid-state Nd:glass laser which operates near 1.05μ , which can be made very small, rugged, reliable, long-lived and inexpensive. Laser radiation near 1μ can readily be transmitted through durable and inexpensive fused silica optical fibers over great distances with negligible loss. This laser wavelength is also readily transmitted through sapphire breech window material. The Nd:glass laser has been extensively used as an ignition source within the LIGHT Program as a result of these attributes.

The aforementioned discussion summarizes key issues which have been considered in the ongoing development of a laser-based ignition system for large caliber guns. The bulk of this paper describes the experimental research and testing in the development of a laser-based ignition system for the Advanced Field Artillery System (AFAS) 155 mm howitzer (Unicharge and Liquid Propellant (LP)) and the Advanced Tank Cannon System (ATACS) 140 mm tank gun round. The goal of the Unicharge laser ignition program at present is to eliminate the current M82 primer from the gun breech and to ignite the blackpowder igniter pad with a laser with no modification to Unicharge. The focus of the LP laser ignition program is to replace or augment the electrical staged ignition system. The long-term future goal is to eliminate all igniter material from the ignition train and to directly ignite the propelling charge with a laser at multiple points to achieve isochronic ignition. The immediate goal of the ATACS laser ignition program is to replace the primer with an optical feed

to couple laser energy into two blackpowder pads within the two piece ammunition. The long-term future goal is to similarly eliminate all primer and igniter material from the round.

EXPERIMENTAL

There are numerous experimental set-ups and configurations of apparatus that are currently being employed in this work. Figure 1 depicts a schematic of a typical laboratory diagnostic set-up. There are numerous diagnostic lasers which are available to probe ignition and combustion events in the laboratory such as a Nd:YAG-Dye laser system which produces tunable laser light. Two high-energy Laser Photonics Nd:glass lasers serve as ignition sources. These lasers are variable energy (up to 30 Joules laser energy/pulse) and can generate pulse widths from 100 μ s to 10 ms. Diagnostics equipment includes optical multichannel analyzers, spectrometers, pressure sensors, digital scopes and other high speed image processing equipment. The laser beam could be focused into an optical fiber bundle with a 9-way split which gave ca. 3 J laser energy at the end of each SMA connector.

The propellant grains were mounted on a high-precision motion stage (a stack of three Daedal series 100000 linear micropositioners and one Daedal series 20000, 5 in rotary table) with four degrees of control (X,Y,Z, Θ). The translational stages each provide 4 inches of travel with a translational accuracy (straight and positional) of $\pm 5.0 \times 10^{-5}$ in./in. of travel and bidirectional repeatability of 5.0×10^{-5} in. The rotational stage provides angular repeatability of 0.2 arc min. with an accuracy of 3.0 arc min. Each stage is driven by a stepper motor with microstepping controlled by a Epson Model Plus microcomputer. Time sequencing of the two lasers was accomplished using a high precision (± 10 picoseconds) digital delay generator (Stanford Research Systems, model #DG 535) which was triggered with the amplified signal from a high-speed pin-photodiode.

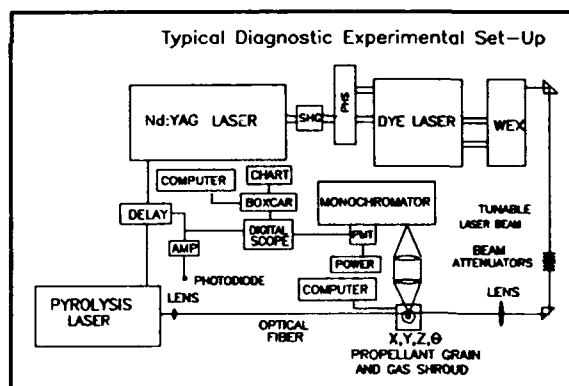


Figure 1. Laboratory Experimental Set-Up.

Small scale testing of the ignition system was performed in an indoor range where larger samples could be studied. Full scale testing was performed at the Interior Ballistics Division Large Caliber Gun Testing Facility.

RESULTS AND DISCUSSION

Blackpowder Ignition (Unicharge)

Interior ballistics (IB) calculations suggest that isochronic laser ignition of multicomponent ammunition can lead to substantial benefits. Localized ignition (non-isochronic), which leads to non-uniform flamespread, may produce pressure differentials (axial pressure waves) which can lead to gun failure. A laser-based ignition system for Unicharge also has the potential to solve stand-off problems in low zones and simplify the ignition train through the elimination of the primer which is inherently difficult to insert and extract using an autoloader system. A conceptual drawing of a Unicharge laser-based igniter train is shown in Figure 2 which requires no alteration to existing Unicharges. The laser beam must pass through several interfaces to be effective:

a sapphire window in the breech, a thin mylar window covering, and a mesh bag which contains blackpowder. The Nd:glass laser beam transmits nearly 100% through the sapphire window. The laser beam must also pass through the mylar window with low loss. Detailed experimentation has shown that regardless of the laser pulse duration (2-10 ms), 70% of the laser was transmitted through the mylar. It was also found that the weaving (mesh) of the bag material was sufficiently loose to readily allow for laser transmission through the bag. A detailed investigation of the laser parameters required to ignite blackpowder in bags through sapphire and mylar windows was performed. A parametric investigation of the laser pulse duration, laser energy and ignition delay on the ignition of blackpowder revealed several reproducible trends. The criterion for ignition was a single laser pulse which resulted in sustained ignition and complete combustion which consumed the entire sample. If the sample did not ignite with a single laser pulse then it was discarded and replaced with an

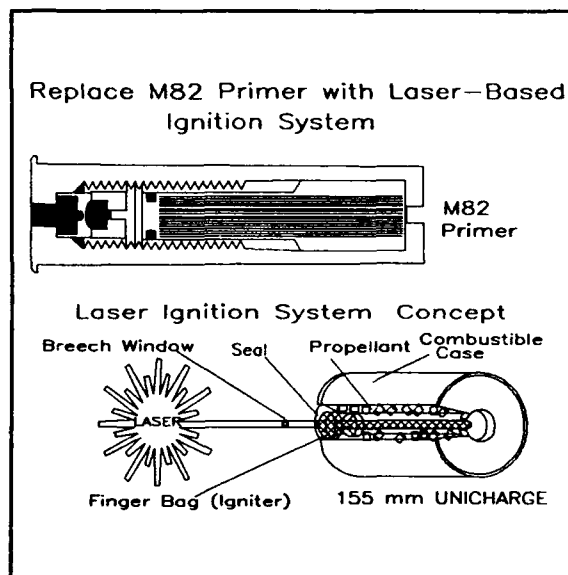


Figure 2. M-82 Primer and Unicharge Laser Ignition Concept.

identical sample. A second laser shot into a previously irradiated sample showed that the first laser pyrolyzed the material which produced new chemical products with reactivities that differed from the original sample. This resulted in a lower ignition threshold. Figure 3 is a schematic which illustrates the trends that were observed in blackpowder ignition. At constant energy of 2 J/pulse, the ignition delay increased with increasing laser pulse widths, but the ignition energy threshold decreased with increasing laser pulse length, i.e., longer laser pulses required less energy for ignition than shorter laser pulses. Regardless of the laser pulse duration, for a given laser pulse; increasing the energy reduced the ignition delay. Although longer laser pulses had a lower energy threshold than shorter pulses, the increased ignition delay time is due to conductive losses early on in time. These results are presented in Figure 4 (a-d).

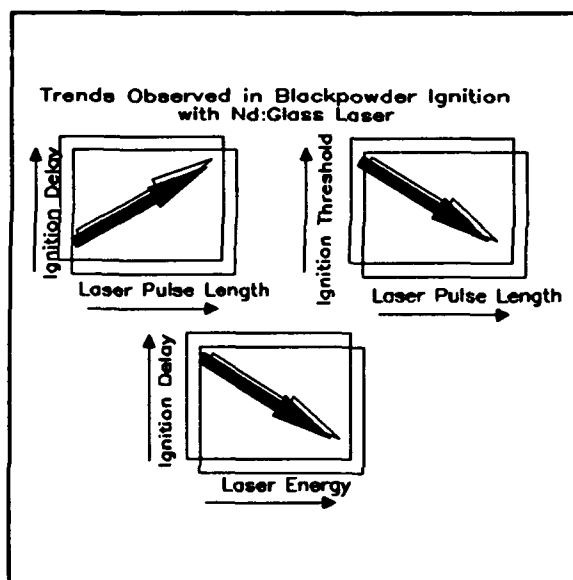


Figure 3. Trends Observed in Blackpowder Ignition.

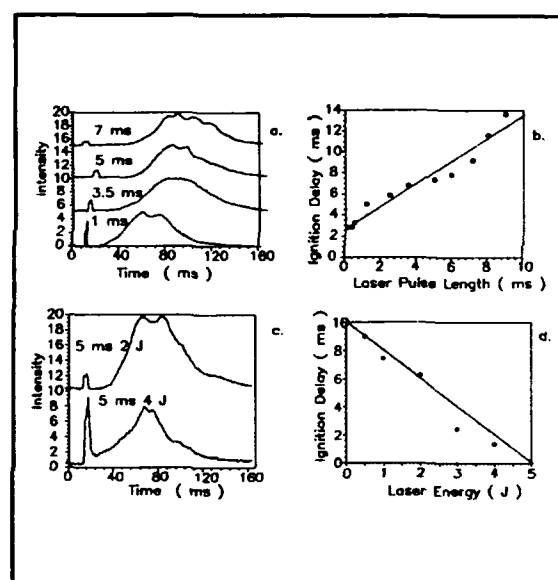


Figure 4. Experimental and Plotted Data for Laser Ignition of Blackpowder.

A key feature of this work which is readily apparent is that although at constant energy longer pulses have a longer ignition delay, simply increasing the energy results in minimal ignition delays. Therefore a plausible method of minimizing ignition delay times for blackpowder is to maximize energy input. Laser pulse lengths 1 ms or greater which deliver several joules of energy will reliably ignite blackpowder charges with a safety factor about five. Extensive laser ignition tests were also performed in an indoor range using large samples of blackpowder (28 grams, 1 ounce) similar to

BARROWS, FORCH, BEYER, COHEN, MIZIOLEK, NEWBERRY

the blackpowder charge contained in a Unicharge increment. It was found that these charges could be readily ignited using the same parameters as the smaller samples. A series of experiments using an optical fiber network for multipoint ignition of blackpowder were performed. Time-intensity traces resulting from the ignition of six one-ounce blackpowder igniters were detected by a single photodiode and demonstrated the near simultaneous ignition of all igniters.

Pressure-time traces from the ignition of blackpowder samples which are confined show that within a few ms after the laser has fired there is a rapid increase in pressure from the gassification process occurring during the blackpowder combustion. These experiments indicate that hot gases required for the ignition of the propellant bed are produced very quickly and can facilitate the spread of the ignition stimulus into the charge.

Laser ignition of blackpowder substitutes such as ball powder and clean burning igniter (CBI) has also been successfully demonstrated. These materials produced less residue than blackpowder, however their ignition and burning rates are pressure sensitive.

Blackpowder Ignition (ATACS)

A laser-based ignition system is also being developed for the ATACS 140 mm two-piece tank round. Interfaces which inhibit rapid and reliable flamespread within the propellant bed can lead to localized ignition which in turn may produce pressure differentials between the charge and projectile base. Interior ballistics calculations performed at BRL have shown that simultaneous laser ignition of the two-piece tank round can enhance flamespreading characteristics and minimize the probability of gun failure. The ATACS round, unlike an artillery charge, utilizes a stub-case which makes a seal to the breech. BRL has developed an ignition concept for ATACS which is presented in Figure 5a. The gun breech contains a sapphire window through which the laser beam is transmitted, however, in addition, the stub case also contains a window. Combustion products may contaminate the stubcase window, but the breech window remains protected from this environment. The next ATACS round which is loaded contains a fresh window. An optical fiber which is contained within the first component of the two-piece ammunition delivers a portion of the laser energy to a blackpowder igniter in the forward ammunition component. The optical fiber in the rear component can easily be contained in an igniter tube or combustible case which will facilitate loading of the propellant. BRL also proposes the use of tapered optical fibers to facilitate the transfer of the laser beam from the breech into the optical fiber contained in the rear component. Optical fiber networks can also be distributed

within the charge to achieve multipoint ignition or to accommodate complex projectile geometries which may extend into the rear portion of the ammunition. Extensive experiments were performed utilizing this ignition concept for the simultaneous ignition of two blackpowder igniters for ATACS. Essentially simultaneous ignition of both igniters was achieved using this system.

Three-point laser ignition of a two-piece tank round configuration in a full scale ballistics simulator has been demonstrated. A single laser pulse distributed through three optical fibers into the rear, center and booster of a center core igniter successfully ignited both ammunition components, Figure 5b.

Direct Propellant Ignition

The overall goal of the LIGHT Program, as previously mentioned, is to eliminate all primer and igniter material from the ignition train in large caliber gun systems. Direct laser-based ignition of a series of propellants using the Nd:glass laser was investigated. These include JA2, M30, LKL, LOVA (XM39, (M43)), HMX1 and CL-20 which were all ignited with a single laser pulse. An important consideration is the coupling of the laser energy into the propellant at the surface.¹⁰ Coatings on the propellant, such as graphite greatly enhance the absorption of laser energy at 1.05μ . In general, all propellants which were graphite coated required less laser energy for ignition. Ignition is also enhanced when graphite is dispersed within the propellant formulation. An additional important consideration is the laser pulse duration. Longer laser pulses on the order of 3-10 ms, 5-10 Joules successfully ignited propellant samples in ambient air. During the initial stages of ignition in a gun system, the propellant is initially at ambient temperature and pressure, however confinement allows for the subsequent pressure build-up which may serve to accelerate burn rate and flame spread throughout the bed. These types of experiments are therefore representative of the early stages of direct laser ignition of the propellant. Unconfined LOVA propellants were very difficult

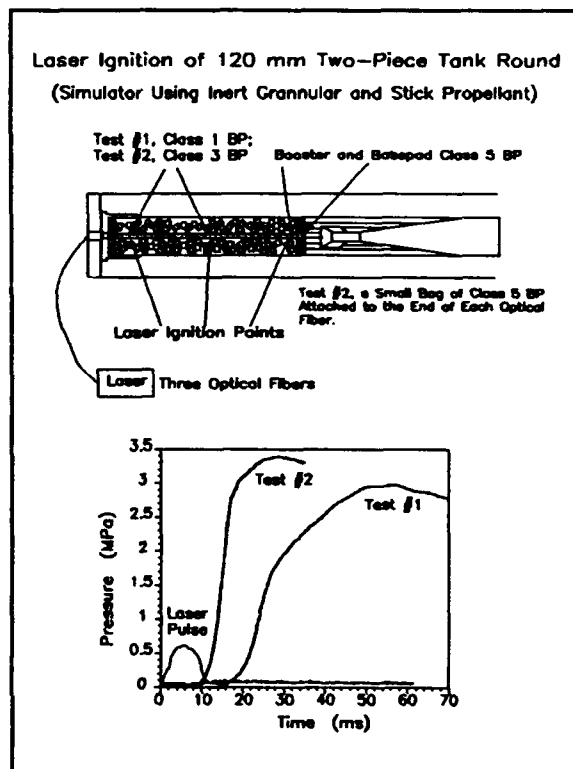


Figure 5. ATACS Laser Ignition Simulator Configuration and Pressure-Time Ignition Plots.

to ignite under ambient conditions. Confined LOVA samples (M43) were successfully ignited with a single laser pulse (10 ms, 8 Joules). These experiments clearly demonstrate feasibility of direct propellant ignition in large caliber guns. Experiments were also performed on early generation LOVA propellants. In particular, the formulations which used a Hexa-terminated polybutadiene (HTPB) binder could be ignited with as little laser energy as 0.5 Joules using a 2 ms laser pulse. Laser radiation at 1.05μ is readily absorbed deep within the propellant which facilitated the ignition.

Representative data from the single laser pulse of JA2 propellant has been investigated. Ignition delay and threshold is sensitive to laser pulse duration, percent surface irradiated and percent oxygen concentration in the ambient atmosphere, Figure 6 (a-c).

Liquid Propellant Ignition

Liquid propellant has been selected as the propulsion system for the AFAS 155 mm large caliber gun. The current ignition system for the LP gun is a staged electrical igniter. Discharge of 45 Joules of electrical energy across an electrode spark-gap ignites a small 2 cc confined volume of LP. The flame propagates into a 20 cc volume then finally into a third 200 stage. The hot-burning LP is injected into the gun chamber wherein up to 17 L of LP is ignited (full charge for high zones). Apparently there are problems associated with flamespreading within the staged ignition and survivability problems with the electrodes for repeated firings. BRL is presently investigating the development of a laser-based ignition system for the LP gun. A laser-based ignition system may allow

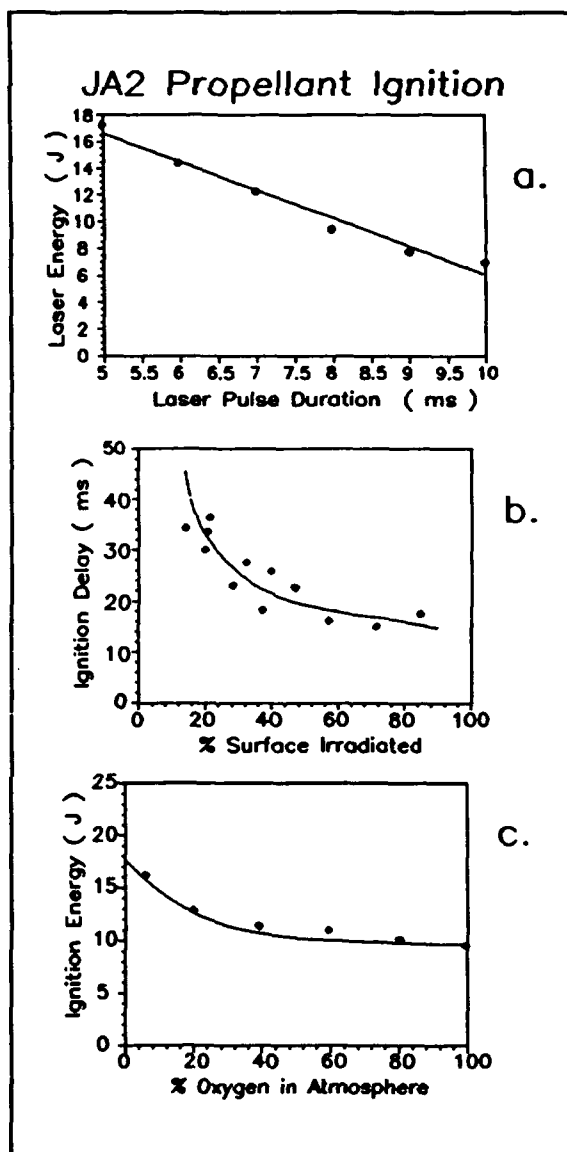


Figure 6. Plots of Trends Observed in the Laser Ignition of JA-2 Propellant as a Function of Laser Parameters.

for one or more stages in the current igniter to be eliminated. A laser-based ignition system could replace existing electrodes in the current igniter or could be used in conjunction with an electric igniter to facilitate flamespread throughout the staged ignition system. A conceptual illustration of this approach is shown in Figure 7a. Laser energy could be coupled into multiple points within the igniter using an optical fiber network. Sapphire windows and ignition points can be located at various points as required. Contamination of the windows may not present the same problem as encountered in solid propellant gun environments.

A series of experiments on LP ignition using pulsed laser systems have been performed. It was found, using a focused-pulsed Nd:glass laser that suspended droplets of LP can be ignited into transient combustion, however sustain combustion is difficult to achieve without confinement. The absorption cross section of LP is much larger in the uv. Focused nanosecond uv pulses at 193 nm ignite the droplets, but combustion was not sustained under

ambient conditions. Additives, such as graphite suspensions and IR dyes were doped into LP samples, however confinement was required for sustained ignition. It is interesting to note that LP samples that did not ignite with a first laser pulse could be ignited or reignited with much less laser energy with a second pulse, Figure 7b. This phenomenon has been attributed to laser decomposition of the LP (pyrolysis) to produce chemically different and more reactive products. These products have not yet been identified.

Double-Laser Pulse Ignition

Experiments aimed at utilizing multiple pulses with a single laser and double pulses utilizing two-laser systems are being investigated for ignition solid propellants¹¹⁻¹² and LP. This concept involves the pyrolysis of the sample to produce

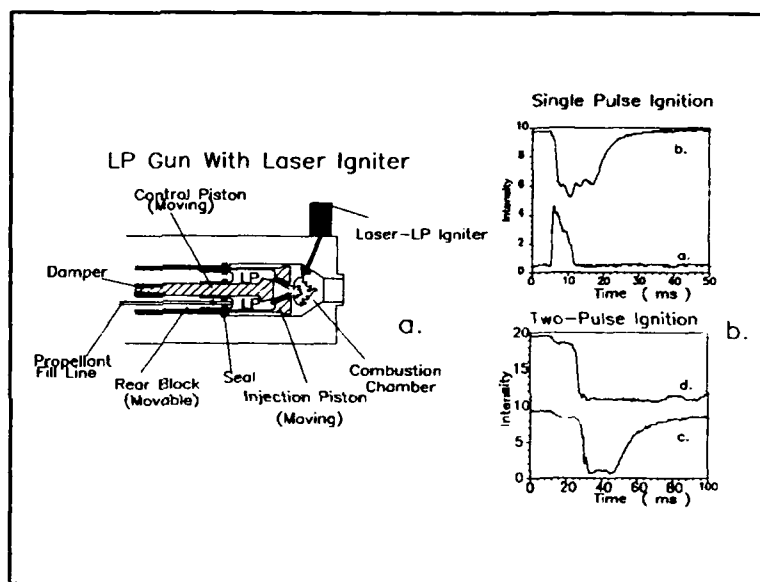


Figure 7. Laser-Based LP Igniter and Ignition Data.

intermediates, which can then be ignited by a second laser pulse from a different wavelength laser, Figure 8a. Initial experiments on the pyrolysis of JA2, for example, have identified the presence of both atomic oxygen and hydrogen. A probe laser system which is tuned to resonant multiphoton excitation transitions of these atoms^{13,14} shows the formation of laser produced microplasmas, similar to those observed in reactive gases, Figure 8b. When the probe beam is tuned to the multiphoton absorption transitions of these radicals, free electrons are produced from multiphoton ionization. Interaction of the intense laser field with the electron field results in non-linear amplification process which generate microplasmas. These microplasmas have sufficient intensity to ignite the pyrolysis gases, but have insufficient thermal feedback to the bulk material to achieve ignition of the solid. Pyrolysis gas diffusion away from the propellant surface may be sufficient to inhibit the ignition. Furthermore, the second probe laser energy is only ca. 1 mJ. This energy range has been shown to be sufficient to ignite reactive gases, but is not sufficient to cause ignition of the bulk solid. Similar types of experiments utilizing an excimer laser^{15,16} in a pump-probe configuration which can deliver more than a factor of 100 times more laser energy, will be investigated.

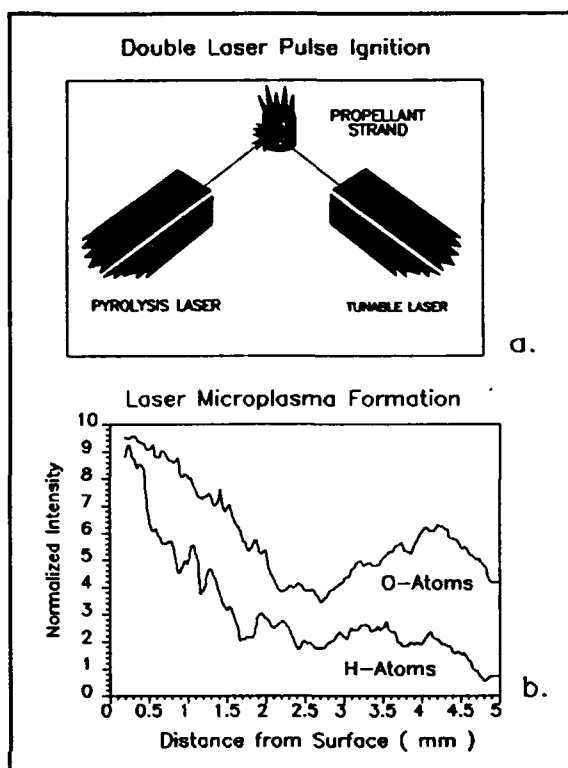


Figure 8. Double Pulse Laser Ignition Concept and Profiles for Laser Produced Microplasmas Through Multiphoton Excitation of O and H Atoms.

CONCLUSION

The development of laser-based ignition systems for large caliber guns has the potential to solve problems associated with reliable and reproducible flamespreading characteristics within propellant beds. The recent advances in gun propulsion systems which utilize multicomponent ammunition, autoloading devices for projectiles, charges and primers place new constraints on the ignition train. In addition, insensitive munition requirements for future gun systems may require alternate or non-conventional ignition sources to be implemented. As a result, laser-based ignition

BARROWS, FORCH, BEYER, COHEN, MIZIOLEK, NEWBERRY

systems may prove to be the only viable initiation source for these munitions. Laser energy distributed through optical fibers embedded in a propellant bed can not only insure simultaneous ignition of the charge, but also reduce overall system vulnerability from the elimination of all primer and igniter material from the munition. Recent advances in electronics and engineering technology have produced small, high-energy laser systems which are suitable as igniters. Extremely small high-power laser diodes are being developed which can easily initiate high explosives and similar energetic materials. These types of devices may prove to be the lasers of choice in future gun systems as new technology developments increase energy output suitable for large caliber gun initiation. New developments in optical fiber material may produce energetic and consumable fibers which leave no residue in gun systems and in addition, enhance ignition. BRL anticipates the development of a laser ignition model of a predictive nature to further understand the chemistry and physics of laser ignition and design igniters with specified characteristics.

ACKNOWLEDGEMENTS

This work was supported by the Ballistic Research Laboratory (BRL), the Program Manager-Advanced Field Artillery System (PM-AFAS) and the Program Manager-Tank Main Armaments System (PM-TMAS). Purchase of data acquisition and processing equipment through the Productivity Capitol Investment Program (PCIP) administered by David Ellis is appreciated.

REFERENCES

1. L. DeLuca, T.J. Ohlemiller, L.H. Caveny and M. Summerfield, *AIAA Journal*, Vol. 14 p. 1111, 1976, and references therein.
2. B.E. Forch and A.W. Miziolek, *Opt. Lett.*, Vol. 11, p. 129, 1986.
3. B.E. Forch and A.W. Miziolek, *Comb. Sci. and Tech.*, Vol. 52, p. 151, 1987.
4. B.E. Forch and A.W. Miziolek, *Comb. and Flame*, Vol. 85, p. 254, 1991.
5. B.E. Forch and A.W. Miziolek, BRL-TR-2740, June 1986.
6. B.E. Forch and A.W. Miziolek, BRL-TR-2809, June 1987.
7. B.E. Forch and A.W. Miziolek, BRL-TR-3295, December 1991.
8. B.E. Forch, J.B. Morris and A.W. Miziolek, Invited Book Chapter, "*Laser Techniques in Luminescence Spectroscopy*", STP 1066, ed. T. Vo-Dinh and E. Eastwood (1990).
9. C.N. Bowden, G.G. Cook and P.S. Henning, Unpublished Royal Armaments Research and Development Establishment (RARDE), 1985.
10. R.A. Pesce-Rodriguez and R.A. Fifer, *Appl. Spectrosc.*, Vol. 45, p. 417, 1991.
11. B.E. Forch and A.W. Barrows, "Detection of Atomic Oxygen and Atomic Hydrogen in Solid Propellant Combustion Using Multiphoton Laser-Induced Fluorescence", submitted to *Comb. and Flame*.
12. B.E. Forch, A.W. Barrows, R.A. Beyer, A.W. Miziolek and J.E. Newberry, "Laser Ignition of Energetic Solid Materials", accepted, *The Journal of Heat and Mass Transfer*.
13. M. Aldén, M. Schawlow, S. Svanberg, W. Wendt and P.-L. Zhang, *Opt. Lett.*, Vol. 9, p. 211, 1984.
14. M. Aldén, H. Edner, P. Grafstrom and S. Svanberg *Opt. Comm.*, Vol. 42, pp. 244, 1982.
15. C.N. Merrow, and B.E. Forch, *J. Chem. Phys.*, Vol. 93, p. 4791, 1990.
16. B.E. Forch and C.N. Merrow, *J. Chem. Phys.*, Vol. 95, p. 3252, 1991.

Correlation of Ballistic Performance to Shear Instability
Studies in High Strength Steels

John H. Beatty*, Dr.,
Morris Azrin, Dr.

US Army Materials Technology Laboratory, SLCMT-EMM
Watertown, MA 02172-0001

INTRODUCTION: Steel is the most commonly used armor material because it is economical and can be easily fabricated. It often serves a dual role as both a structural member and protective armor in combat vehicles. For these reasons, steel will continue to be an important component of many advanced armor systems for the foreseeable future.

The two properties of steel that most affect its ability to resist penetration by projectiles are the hardness and cleanliness of the steel. For many years, the general relationship between increased resistance to penetration at higher hardnesses has been known,^{1,2} with no other single metallurgical factor as important. However, reduced resistance due to shattering of steels harder than 52-54 on the Rockwell C scale (HRC) drastically limits their usage above this hardness range. The cleanliness of the steel also affects the ballistic response,³ as cleaner steels have fewer impurity inclusions. Inclusions degrade both the resistance to penetration and shattering resistance of the armor simultaneously. Until recently, the dominant factors of hardness and cleanliness have obscured the influences of many other metallurgical factors on ballistic performance.

The advent of processing techniques such as Vacuum Arc Remelting (VAR) and Electroslag Remelting (ESR) allow steelmakers to achieve greatly reduced impurity levels and inclusion contents, and this in itself has led to improvements in ballistic properties.³ These techniques have also allowed for the examination of very "clean" steels to understand the influences of other microstructural components without the overshadowing presence of the inclusions. Recent studies⁴ in VAR and ESR 4340 steels suggested that grain-refining carbides can play an important role during ballistic penetration, and this study is directed towards obtaining a

fundamental understanding of how and to what degree these carbides alter the ballistic behavior of high strength steels.

BACKGROUND: Microstructures of Armor Steels

Armor steels are used in a wide variety of metallurgical conditions depending upon the particular application. Rolled Homogeneous Armor (RHA)⁵ and High-Hard Armor⁶ are the two most common types, with each being available in different thicknesses and hardnesses.

To obtain optimum ballistic properties, steel should be processed to produce a fully martensitic microstructure. Martensite is a phase that can form in armor steels upon rapid cooling (quenching). Martensite is extremely hard and strong, which is beneficial to ballistic properties. However, in the as-quenched state, the martensite is usually too brittle, and must be tempered to improve the toughness of the steel. The increase in toughness gained through tempering comes at the expense of a lower hardness. For each application, the trade-off between hardness and toughness should be appropriately balanced to provide sufficient structural integrity and maximum ballistic protection.

The resulting microstructure of steel armor after quenching and tempering is quite complex, and a typical example is shown in Figure 1, along with a schematic. The martensite in these steels is arranged in a lath structure, with individual laths ranging from several to tens of microns in width. The laths are found in several orientations or "packets" within each martensite colony. In addition to the lath structure, other important microstructural components are present. Grain boundaries remain from the austenitic phase that preceded martensite formation, and these are referred to as prior austenite grain boundaries. Along these prior austenite grain boundaries, grain-refining carbides (usually between 0.5-0.05 μm in diameter) are found. This microstructural component is the central focus of the present study. Grain refining carbides should not be confused with temper carbides which are precipitated during tempering and are much smaller, needle-shaped precipitates. Finally, impurity inclusions can also be found within conventionally produced steels. These typically range in size from 1 -10 μm in diameter, and these particles are very detrimental to ballistic properties. As previously mentioned, refinement techniques such as VAR or ESR remove most of these impurities. The steel used in this study, VAR 4340, has little, if any, of these inclusions so the effects of other microstructural factors on ballistic properties can be examined.

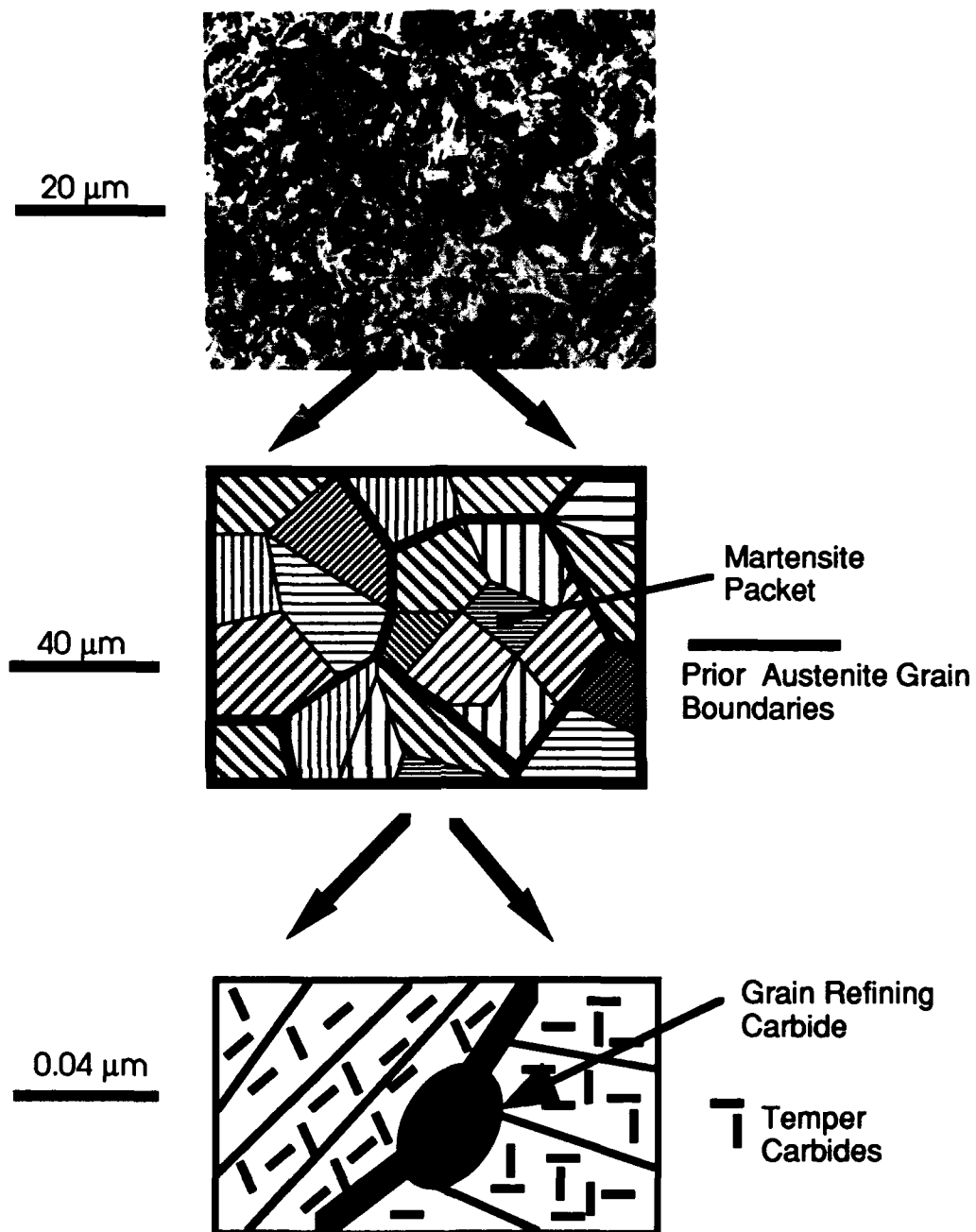


Figure 1. Microstructure of VAR 4340 steel. a) Optical Micrograph showing lath structure. b) Prior austenite grain boundaries and packets of lath martensite. c) Grain-refining carbides are found along prior austenite grain boundaries, temper carbides are found within the laths.

Ballistic Failure of Thin Armor Plate

A shearing mechanism called plugging is often the predominant failure mode during ballistic penetration of thin, high-hardness steel armor plate. Figure 2 illustrates this failure mode. The projectile impacts the armor, and the material surrounding the impact zone starts to undergo plastic shear deformation. Kinetic energy from the projectile is transferred to the plate by the plastic deformation, and the armor in this region begins to increase in temperature. Because the deformation is occurring very rapidly, there is insufficient time for the steel to conduct heat away from the deforming zone so the deformation can be considered adiabatic. Two processes are competing with each other at this stage. The deformed substructure within the material causes stiffening (work hardening). Simultaneously, the increase in temperature created by the rapid energy absorption softens the material. At some point in time, the rate of thermal softening surpasses the rate of work hardening and the material becomes unstable. The load bearing capacity in this area is reduced by the thermal softening, which leads to even more severe deformation, which in turn causes even more thermal softening, and so on. This shear instability causes the localization of the shear strain into a narrow region known as an adiabatic shear band.

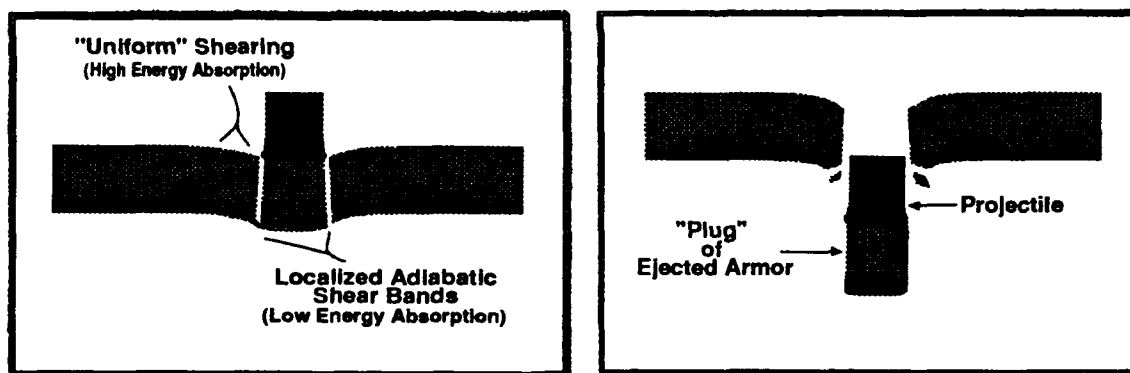


Figure 2. Cross sectional depiction of failure of armor by plugging. The projectile strikes the armor and creates localized shearing. A plug of armor is ejected through the rear face of the plate.

The adiabatic shear band will grow to surround a cylindrical plug of armor. Since the material within the band continues to thermally soften at an increasing rate, little energy is absorbed by the plate from this point on. The

remaining kinetic energy of the projectile is divided between itself and the plug, and both the plug and projectile exit the rear of the armor plate at high velocity, causing damage to any personnel and equipment in their paths. Delaying or preventing the onset of the adiabatic shear banding will allow the armor to absorb more of the projectile's kinetic energy, which in turn can prevent complete projectile penetration of the armor.

GENERAL APPROACH

It is the aim of this study to characterize how microstructural factors affect the formation of adiabatic shear bands. To accomplish this, it is necessary to control the microstructures being evaluated, and to examine their response to shear deformation over a variety of strain rates. An innovative heat treatment has been used to alter the size and distribution of the grain refining carbides, while leaving the rest of the microstructure virtually unchanged. Laboratory shear tests have been performed over a variety of stress-states and strain-rates. Finally, the ballistic performance under plugging conditions has also been evaluated. Performing these shear instability tests on several carefully controlled armor microstructures produces a more complete understanding of the plugging process. This should eventually lead to improved armor alloys and armor processing techniques.

EXPERIMENTAL PROCEDURES

The alloy used in the current study is a VAR 4340 steel. Its composition and the four heat treatment schedules utilized are given in Tables 1 and 2, respectively. Double austenitization treatments were used to produce four different microstructures, all with a hardness of 52 HRC. A unique distribution of grain refining carbides is produced during the initial austenitization because the carbide dissolution and coarsening rates vary at each temperature. The second short austenitization is performed to obtain a uniform prior austenite grain size. The time of the second austenitization is not sufficient to significantly change the carbide distribution. Thus we are able to produce four microstructures with equivalent hardnesses, similar prior austenite grain sizes, but with unique distributions of grain refining carbides. These microstructures have been examined in our earlier studies.^{7,8} Microstructural characterization included optical microscopy, scanning electron microscopy, and transmission electron microscopy.

Ballistic testing was also performed on each microstructure. A standard V50 ballistic performance evaluation was made using two 12"X12"X3/8" plates of each microstructure with 0.50 caliber APM2 projectiles fired at 0° obliquity. The V50 is defined as the velocity at which a given

Table 1								
Alloy Composition, wt %								
C	Mn	P	S	Si	Cr	Ni	Mo	Cu
0.42	0.46	0.009	0.001	0.28	0.89	1.74	0.21	0.19

Table 2				
Heat Treatment Schedule				
1st Austenitization Temperature, °C 1hr	oil quench	2nd Aust. Temp., °C 15 min.	oil quench	Temper, 2hrs. °C
845	X	845	X	200
925	X	845	X	200
1010	X	845	X	200
1090	X	845	X	200

projectile has a 50% probability of creating a "complete" penetration of the armor plate. A complete penetration is one in which some fragments of the either the projectile or the armor plate perforate a thin aluminum witness plate placed behind the target armor. Commonly, as is the case for the current tests, the measured V50 is the average velocity of six different shots (3 partial and 3 complete penetrations) that fall within a narrow specified velocity range of 125 ft/s.

The ballistic data will be compared to shear instability tests over a variety of strain rates. Shear instability tests were completed by the authors and other co-workers as part of a comprehensive study⁷⁻⁹, and have been reported in detail previously. A brief description of each technique follows.

Quasi-static shear tests were performed using thin-wall torsion samples and double linear-shear^{7,9} specimens. The double -linear shear specimen configuration is given in Figure 3. For each microstructure, a shear stress-displacement curve was obtained. Additionally, scribe lines through the gage length of each specimen were used to obtain a strain profile for each sample by measuring the angle of deflection of the scribe line. Each test was interrupted once the load decreased approximately 2% from the maximum. The maximum uniform shear strain was found using the strain profiles, and is referred to as the shear instability strain, γ_i .

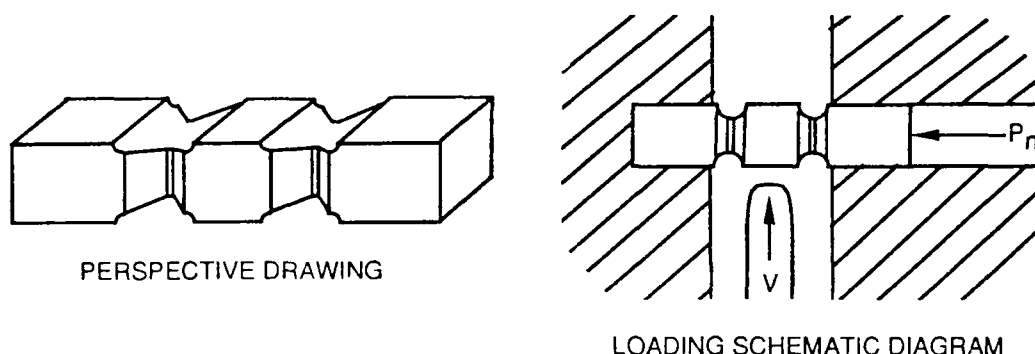


Figure 3. Schematic of the double linear shear specimen and its loading configuration. From references 7 and 9. For both quasi-static and dynamic tests. An axial load, P_n , can be imposed. "V" indicates the velocity of the plunger or pendulum causing shear. Only quasi-static data are reported in this paper.

High strain rate shear behavior was characterized using a hat-shaped specimen (developed by Meyer et al.^{10,11}) in a split-Hopkinson compression bar (Figure 4). These experiments were part of a joint program with the University of California, San Diego.^{8,12} This cylindrical sample creates adiabatic shear bands, while allowing experimental control of the strain-rates and/or displacements induced. The reflected and transmitted elastic waves in the bar allow for the measurement of stress-displacement curves for each microstructure. From this information, the energy required to initiate unstable shearing can be determined.

RESULTS AND DISCUSSION

The quasi-static shear behavior data is given in Figure 5. The greatest resistance to unstable shearing is found for the 925°C austenitizing temperature. Cowie⁸ showed that the instability strain was proportional to the carbide particle size / average particle separation distance ratio. This behavior is consistent with theories of destabilization by microvoid nucleation softening. The particles themselves cause local tensile stress fields when the surrounding material undergoes pure shear. These local tensile fields are responsible for nucleating voids. Both the particle size and the particle separation are related to this instability. This mathematical

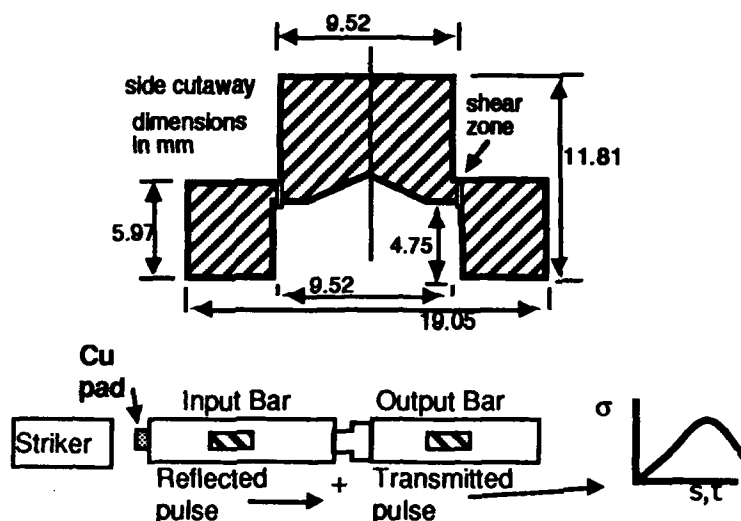


Figure 4. Side cutaway of Hat specimen with split Hopkinson apparatus.

functionality suggests that the interaction of microvoids that formed on particles of average size creates the initial softening of the material and localization of the shear strain.

This view of microvoid softening in high strength steel is also supported by the stabilizing influence of hydrostatic compression on shear deformation.⁹ The application of a compressive stress component reduces the local tensile stresses created at the carbides and delays the nucleation of voids. Therefore, the resistance to unstable shearing is aided by the application of compressive stresses at moderate strain rates.

The high strain-rate shear behavior is shown in Figure 6. Since all four microstructures have equivalent shear strengths, the energy absorbed to the point of instability is directly related to the shear instability strain. Therefore, Figures 5 and 6 can be qualitatively compared. The 925°C austenitization temperature still produced the best results, but the dependance on the carbide size and distribution is unclear. At the present time, there is no explanation for the observed relationship at high strain-rates.

Microstructural characterization studies of specimens deformed at high strain rates may aid in the determination of the relationship between carbides and adiabatic shear in high strength steels⁸. Figure 7 shows a transmission electron micrograph of material taken from the center of a

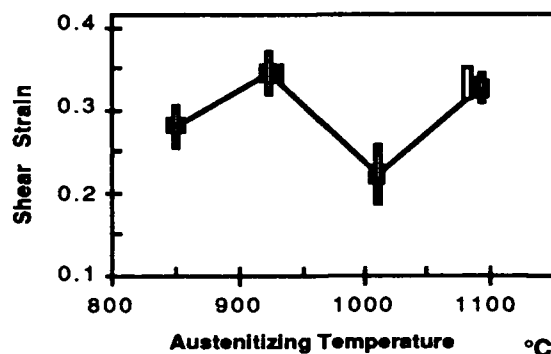


Figure 5. Shear instability strain as a function of the first austenitizing temperature for VAR 4340. The solid squares indicate the instability strain, the vertical bars give the approximate error. A two-step austenitization was used to produce four microstructures all at Rc 52 with a constant prior austenite grain size.

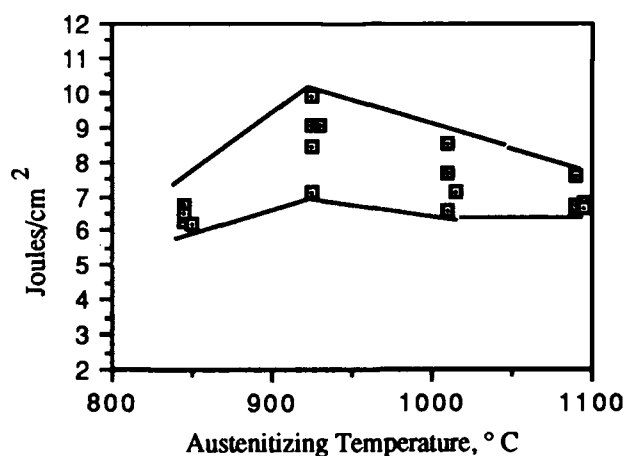


Figure 6. Energy absorbed to the point of instability per unit area sheared versus austenitizing temperature, using the hat specimen in the split-Hopkinson compression bar.

deformed hat specimen. An extremely fine microcrystalline structure is observed, and no carbides can be detected within the shear band. The microstructure is similar in appearance to ball milled metallic powders. It is anticipated that continued analyses of deformed microstructures will give additional insight into the microstructural evolution within the shear band. This should lead to the production of improved microstructures for armor applications.

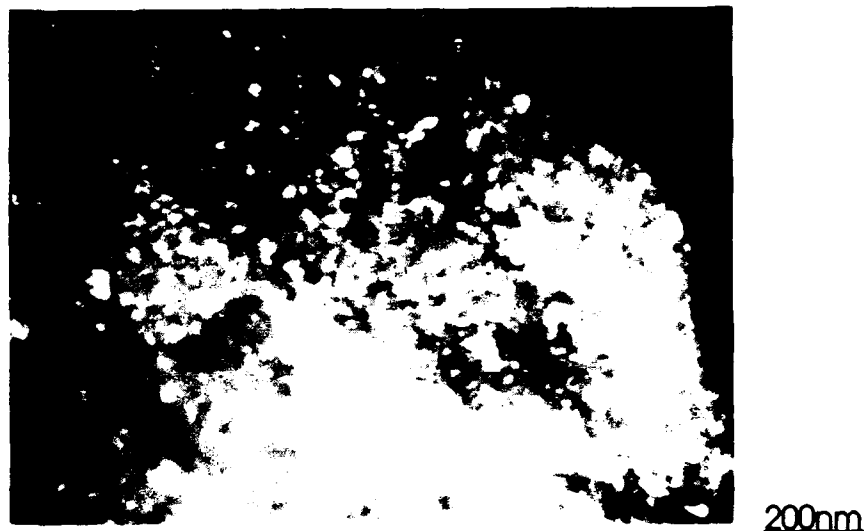


Figure 7. Transmission electron micrograph of material taken from the center of an adiabatic shear band formed in a hat specimen. The structure is microcrystalline, with a crystallite size of 8 - 20nm.

Ballistic V_{50} results for all four microstructures are given in Figure 8. The V_{50} performance closely resembles the results of the hat test. Again, the 925°C austenitizing temperature produced the best microstructure, and a single peak in resistance to unstable shear was observed. These results and previous work by Meyer et al.^{10,11} suggest that the hat test can serve as a controlled laboratory analogue to ballistic testing of thin armor plate if plugging failures predominate. This provides us with an excellent experimental method for studying microstructural effects on armor performance. It is also important to recognize that the carbides have significantly affected the ballistic performance. Though the mechanisms remain unknown, these small components of the microstructure demonstrate that proper microstructural control must be maintained in order

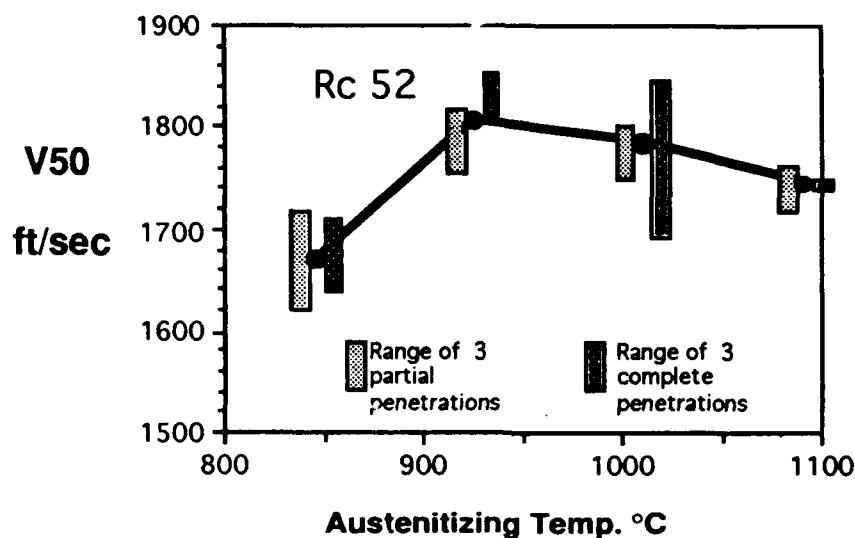


Figure 8. Ballistic results of all four microstructures. The ballistic results are in good agreement with the hat test data given in Figure 6.

to optimize ballistic properties. These same results should help to develop new, unique microstructures that can lead to improved armor materials for future Army applications.

CONCLUSIONS

- Grain refining carbides significantly effect the shear behavior of high strength VAR 4340 steel.
- At low to moderate strain rates, a microvoid softening model can explain the observed experimental results.
- High strain-rate hat tests and ballistic evaluations show good agreement, so the hat test can be used to study armor plugging.
- The tailoring of common and unconventional microstructures should be explored to maximize the survivability of future Army systems.

BEATTY and AZRIN

REFERENCES:

1. S.J. Manganello and K.H. Abbott, Journal of Materials, JMSLA, vol. 7, no. 2, pp.231-239, June 1972.
2. J.F. Mescall and H. Rodgers, "The Role of Shear Instability in Ballistic Penetration" , Proceedings of the 34th Sagamore Army Materials Research Conference, pp. 287-314, August 30-September 3, Lake George, NY, 1987.
3. C.F. Hickey, T.S. Thomas, and A.A. Anctil, "Comparison of Ballistic Performance of a Split Heat of ESR and VAR 4340 Steel," US Government Report, US Army Materials and Mechanics Research Center, Watertown, MA, AMMRC TR 85-20, July 1985.
4. E.T. Myskowski and A.A. Anctil, "Effect of Microstructure on Ballistic Performance of ESR 4340 Steel", US Government Report, US Army Materials Technology Laboratory, Watertown, MA, MTL TR 89-73, March 1989.
5. MIL STND 12560G, Plate, Wrought, Homogeneous, 24 JAN 85.
6. MIL STND 46100C, Plate, Wrought, High Hardness, 20 OCT 86.
7. J.G. Cowie, "The Influence of Second Phase Dispersions on Shear Instability and Fracture Toughness of Ultrahigh Strength 4340 Steel" US Government Report, US Army Materials Technology Laboratory, Watertown, MA, MTL TR 89-20, March 1989.
8. J.H. Beatty, L.W. Meyer, M.A. Meyers, and S. Nemat-Nasser, "Formation of Controlled Adiabatic Shear Bands in AISI 4340 High Strength Steel" US Government Report, US Army Materials Technology Laboratory, Watertown, MA, MTL TR 90-54, November 1990. Also published in *Shock Waves and High Strain-Rate Phenomena in Materials*, editors Meyers, Murr, Staudhammer, published by Marcel Dekker, 1991, in press.

BEATTY and AZRIN

9. J.G. Cowie, M. Azrin, and G.B. Olson, "Microvoid Formation during Shear Deformation of Ultrahigh Strength Steels", *Metall. Trans. A.*, **20A**, pp. 143-153, January, 1989
10. K-H Hartman, H.D. Kunze, and L.W. Meyer, "Metallurgical Effects on Impact Loaded Materials," in *Shock Waves and High-Strain-Rate Phenomena in Metals*, Meyers and Murr, eds, pub Plenum Press, New York, 1981.
11. L.W. Meyer, S. Manwaring, "Critical Adiabatic Shear Strength of Low Alloyed Steel Under Compressive Loading," in *Metallurgical Applications of Shock-Wave and High-Strain-Rate Phenomena*, editors Murr, Staudhammer, and Meyers, pub Marcel Dekker, Inc, New York, 1986.
12. J.H. Beatty, Y-F Li, M.A. Meyers, and S. Nemat-Nasser, "Adiabatic Shear Banding in High-Strength Alloys," to be published in the Proceedings of the Army Solid Mechanics Conference, Plymouth MA, November, 1991.

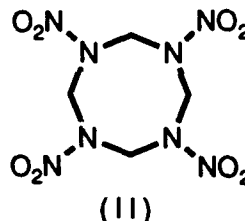
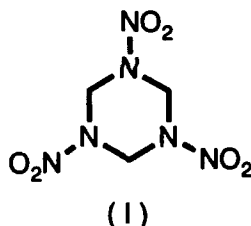
**Thermal Decomposition of RDX and HMX in the Condensed Phase:
Isotope Scrambling and Deuterium Isotope Effect Studies**

Richard Behrens, Jr.
Combustion Research Facility
Sandia National Laboratories
Livermore, CA 94551

Suryanarayana Bulusu *
Energetic Materials Division
U. S. Army Armament Research, Development, and Engineering Center
Picatinny Arsenal, NJ 07806-5000

Introduction

The cyclic nitramines 1,3,5-trinitro-hexahydro-s-triazine RDX (I) and octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine HMX (II)

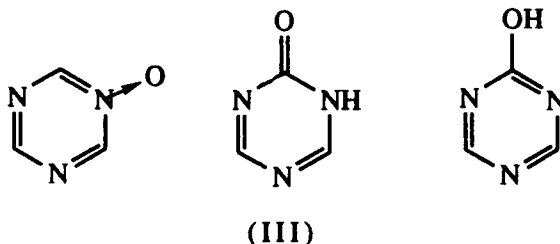


are energetic ingredients that are used in various propellants and explosives. Understanding the complex physicochemical processes that underlie the combustion of these materials can lead to methods for modifying the propellant and explosive formulations in order to obtain better ignition, combustion, or sensitivity properties. Our work strives to provide a link between the physical properties and molecular structures of different nitramines and their combustive behavior. This requires the understanding of the reaction kinetics and transport processes in both the gas and condensed phases. The thrust of our work is to obtain an understanding of the physical processes and reaction mechanisms that occur in the condensed phase so that the identity and rate of release of the pyrolysis products from the condensed phase can be predicted, as a function of pressure and heating rate, based on the physical properties and molecular conformation of the material.

Reviews¹ of the literature on RDX and HMX have discussed the roles of unimolecular decomposition and autocatalysis on the thermal decomposition of these compounds. Previous work²⁻⁵ has shown that HMX decomposes in the condensed phase and that the identity and rates of release of the pyrolysis products are determined by complex physical processes coupled with several different chemical decomposition mechanisms.^{4,5} The physical pro-

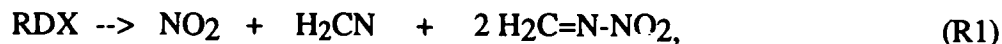
cesses include the formation and containment of the gaseous pyrolysis products in bubbles within the HMX particles and the subsequent release of the gases as escape paths form within the HMX particles. The long containment time of decomposition products within the HMX particles allows secondary reactions between the decomposition products and the HMX.

In contrast to HMX, RDX has a lower melting point (202°C for RDX compared to ~280°C for HMX), so that decomposition occurs primarily in the liquid phase with a faster release of the gaseous products from the sample than for HMX. Although both RDX and HMX have been found^{4,6,7} to form H₂O, HCN, CO, CH₂O, NO, N₂O, amides (e.g., methylformamide), a nonvolatile residue (NVR) and the mononitroso analogues of the parent compound (1-nitroso-3,5,7-trinitro-1,3,5,7-tetrazocine (ONTNTA) from HMX and 1-nitroso-3,5-dinitro-hexahydro-s-triazine (ONDNTA) from RDX), only in experiments with RDX^{6,7,8} have NO₂, and oxy-s-triazine (OST) (III) been observed. The precise structure of OST has not been determined, but it is likely to be one of the following structures:



This paper presents the results of measurements of isotopic scrambling of products from mixtures of isotopic analogues of RDX and of the observed deuterium kinetic isotope effects (DKIE). Together, these two types of results are used to probe the decomposition mechanisms of RDX.

The basic mechanisms that have been considered in the past to describe the decomposition of RDX fall broadly into three categories. These include the least energy pathway⁹ involving the N-N bond cleavage,



the concerted triple scission pathway,



found in infrared multiphoton dissociation (IRMPD) experiments with RDX¹⁰, and the HONO elimination pathway¹¹



Secondary reactions of methylenenitramine, R4 and R5



have been postulated for many years to explain the formation of HCN, CH₂O, and N₂O and have been shown to occur in the IRMPD experiments¹⁰ with RDX.

Recent thermal decomposition studies of RDX^{7,8} have shown that it is also important to account for the formation of OST and ONDNTA in the decomposition mechanism since they may be intermediates from which the final products are formed. To address these issues, mixtures of isotopically labeled compounds are used to track the breaking and formation of chemical bonds that lead to the generation of the thermal decomposition products of RDX. We also use the DKIE to probe the role of carbon-hydrogen bonds in the formation of the different decomposition products. We present results from the previously described simultaneous thermogravimetric modulated beam mass spectrometry (STMBMS) measurements from RDX, RDX-d₆ and mixtures of unlabeled RDX (RDX-ul) with the isotopically labeled analogues, RDX-d₆, RDX-¹³C, and RDX-¹⁵N₆, and those between RDX-¹³C and RDX-¹⁸O. The same measurements are used to determine the DKIE on the pyrolysis products during the course of the decomposition and to determine the amount of isotopic scrambling from different isotopic mixtures of RDX. We report the calculated fraction of isotopic scrambling observed for most of the products formed in experiments with different combinations of the labeled RDX compounds.

Experimental

Instrument description. The STMBMS apparatus and basic data analysis procedure have been summarized in the previous paper⁷ and are described more extensively elsewhere.^{12,13}

Sample preparation. Samples of RDX with different isotopic labels (see Table 1) were synthesized by the general method previously reported elsewhere¹⁴. The isotopic enrichment of each of the isotopomers and the amounts of each one used in a given experiment are given in Table I.

Samples for experiments with mixtures of isotopic analogues of RDX are prepared by mixing equal amounts in acetone solution followed by evaporation of the solvent.

For each mixed-isotope experiment listed in Table I, experiments were run on each of the isotopic analogues separately. The results from these experiments are used to check for the presence of ion signals at m/z values that should originate only from the mixing of isotopes in the scrambling experiments. If there is a measurable amount of ion signal present in the experiments with the separate isotopic analogues, then the results of the separate experiments are used to correct the data from the isotope scrambling experiments.

Table I. Experimental Parameters^a

Compound	Enrichment (%)	Experiment reactant quantity (mg)			
		1	2	3	4
RDX-ul		2.04	2.68	3.35	
RDX-d ₆	97.3	1.98	2.6		
RDX- ¹⁵ N ₆	98.5			3.53	
RDX- ¹³ C	98.1				2.26
RDX- ¹⁸ O	3.7				2.26
Electron energy (eV)	-	18	20	20	20
Orifice dia. (cm)		0.00254	0.00254	0.00254	0.00254
Temp range ^b		198 - 207	198-207	198 -208	198-207

^a All mixtures of isotopic analogues of RDX are recrystallized from an acetone solution.

^b Temperature range over which the decomposition of RDX occurs in °C. Heating rate for all experiments is 0.5°C/min in the temperature range of the decomposition.

Analysis

The data analysis procedures used for identifying the pyrolysis products and determining their rates of formation have been described previously.^{12,13,17} Since this paper examines the extent of isotopic scrambling and the DKIE on the relative rates of formation of the various products, it is sufficient to compare the ratios of the ion signals measured at different *m/z* values for the various products. The structures of the more complex decomposition products discussed in this paper are reasonable assumed structures based on the mass spectra of different isotopic analogues of RDX.⁷ OST has been investigated previously by tandem mass spectrometry⁸ but a definitive structure could not be assigned.

The relative size of the ion signals at the different *m/z* values associated with an individual decomposition product is dependent on the amount of each isotopically labeled analogue used in the experiment, the fraction of each analogue that participates in forming isotopically scrambled products, the cracking factors in the ion fragmentation process, and the fraction of isotopic enrichment in each isotopically labeled analogue.

Briefly, the amount of each product isotopomer formed from the decomposition of a mixture of *r* moles of isotopomer 1 and *s* moles of isotopomer 2 of RDX is given by

$$r \text{ RDX}_1 + s \text{ RDX}_2 \rightarrow x k r (1 - f) A_{11} + k s (1 - f) A_{22} + k f (x r + s) A_{12} \quad (\text{R6})$$

*A*₁₁ and *A*₂₂ are pyrolysis products formed from RDX₁ and RDX₂ respectively, *A*₁₂ is the set of mixed isotope products, *k* is moles of product *A* formed per mole of RDX₂, *x* is the ratio of moles of product *A* formed from RDX₁ to the moles of product *A* formed from RDX₂

(for deuterium-labeled RDX this corresponds to the DKIE factor), and f is the fraction of the product molecules produced from either RDX₁ or RDX₂ that participate in isotope scrambling.

R6 is then combined with the following expression that accounts for the relative number of ions formed in the mass spectrometer at different m/z values from the product isotopomer

$$S_k^+ = \sum_{i=1}^{\text{\# isotopomers}} F_{i,k} A_i^+ \quad (1)$$

where S_k^+ is the ion signal at $m/z=k$, A_i^+ is the total number of ions formed from the isotopomer A_i with mass m_i , and $F_{i,k}$ is the fraction of the total number of ions formed from A_i that appear at $m/z=k$.

From reaction R6 and Eqn.1 expressions for the amount of ion signal at each different m/z value associated with a decomposition product are formulated and used to construct equations based on the ratios of the ion signals at different m/z values for a given product. These equations are then solved for the values of f and x .

Table II. H₂O Isotopic Scrambling Results

Isotopes	Exp #	m/z	Rel Intensities	Temporal Corr
H/D	2 ^a	18	64±11	Y
		19	100	Y
		20	24±3	Y

^a Results for isotopic scrambling of H₂O from experiment 1 are not included because of interference from extraneous signal at $m/z = 18$.

Results

The isotopic scrambling results for the products H₂O, CH₂O, N₂O, OST and ONDNTA based on experiments outlined in Table 1, are summarized in Tables II to VI while the mixing fractions (crossover), f , and the observed DKIEs, x , for the deuterium labeled products are presented in Table VII.

Figures 1 through 6 show the relative intensities as a function of temperature and the temporal correlations of groups of ion signals at various m/z -values that are used to determine the extent of the DKIE and isotopic scrambling for the decomposition products formed from RDX. Since the heating rates were linear with time, the curves also reflect the temporal correlations between the ions signals. All of the data presented in Figures 1 through 6 clearly show

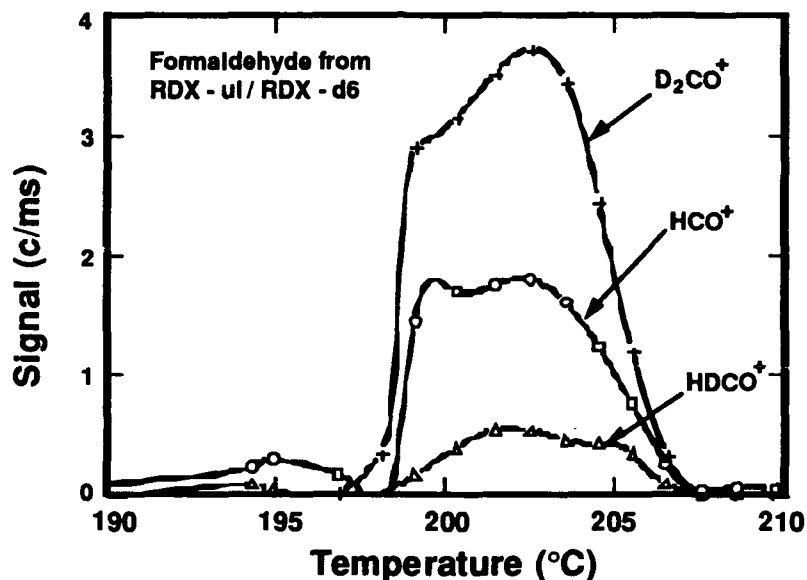


Figure 1. The ion signals from CH_2O and its isotopic analogues formed during the decomposition of a mixture of RDX-ul and RDX-d6 (Experiment 1 in Table I). Between 193 and 210 °C the heating rate of the sample is 0.5 °C/min. The rapid increase in the rate of evolution of CH_2O at 198 °C occurs as the sample melts.

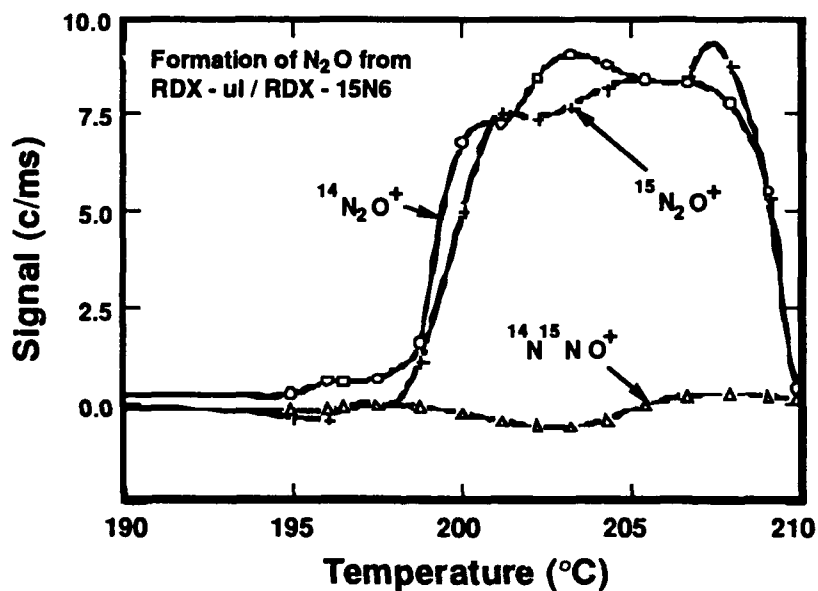


Figure 2. The ion signals from N_2O and its isotopic analogues formed during the decomposition of a mixture of RDX-ul and RDX- $^{15}N_6$ (Experiment 3 in Table I). Between 195 and 210 °C the heating rate of the sample is 0.5 °C/min. The rapid increase in the rate of evolution of N_2O at 199 °C occurs as the sample melts.

Table III. CH₂O Isotopic Scrambling Results

Isotopes	Exp #	m/z	Rel Intensities	Temporal Corr
H/D	1	29	49±3	Y
		31	15±3 ^a	Y
		32	100±3	Y
	2	29	61±5	Y
		31	16±5	Y
		32	100±5	Y
¹³ C/ ¹⁸ O	4	31	100	Y
		32	5.4±0.6	Y
		33	1.4±0.1 ^b	Y

^a Corrected for 31⁺ contribution from pure RDX-ul.

^b Corrected for the 33⁺ contribution observed in a separate experiment with RDX-¹³C. The corrections are made using the CH₂O (¹³C m/z=31) signal for normalization.

that the decomposition products are first observed when the RDX starts to melt at 198°C. As shown in previous paper⁷ decomposition of RDX is significantly inhibited in the solid phase.

H₂O. Scrambling of the hydrogen isotopes is observed in the formation of water from a mixture of RDX-ul and RDX-d₆. Analysis of the data in Table II shows that the fraction of water that participates in mixing is 1.1±0.1, which represents complete mixing, and that the DKIE is 1.5±0.1. Since low electron energies of 18 and 20 eV are used in the mass spectrometer, contributions from OD⁺ are ignored. The results are calculated assuming only the molecular ions, H₂O⁺, HDO⁺, and D₂O⁺ contribute to the measured signal. Since H and D exchange is rapid in water and is likely to occur before detection of the product, the observation of complete H/D scrambling does not necessarily imply that water is produced in a bimolecular reaction between either RDX itself or RDX and a decomposition product.

CH₂O. As indicated in Table VII the fraction of RDX that forms formaldehyde with hydrogen and deuterium exchange is about 7 percent for the experiments. It can be seen from Figure 1 that there is a high degree of temporal correlation between the ion signals from the isotopic analogues of formaldehyde. The observed DKIE for formaldehyde formation is 1.05±0.1, which implies no deuterium kinetic isotope effect. The ion signals used in the analysis are for m/z values of CH₂O⁺, CHO⁺ and their isotopic analogues. Measured ion fragmentation factors for CH₂O and CD₂O were used to calculate the ratio of CHO⁺/CH₂O⁺ and CDO⁺/CD₂O⁺ formed in the ionization process.

Experiment 4 with a mixture of RDX-¹³C and RDX-¹⁸O shows that 45±7 percent (Table VII) of the formaldehyde formed is from mixing between the two differently labeled RDX

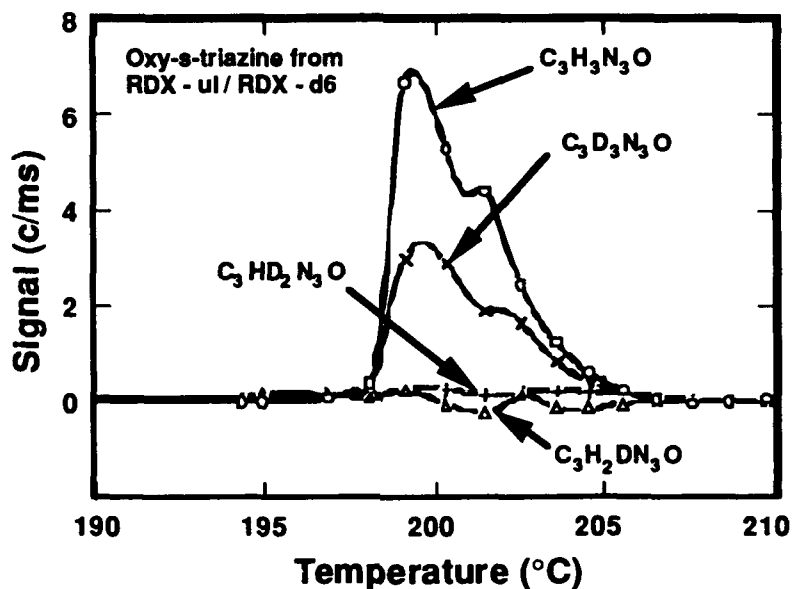


Figure 3. The ion signals observed from oxy-s-triazine and its isotopic analogues formed in the decomposition of a mixture of RDX-ul and RDX-d6 (Experiment 1 in Table I). Heating rate of the sample is 0.5 °C/min. The rapid increase in the rate of evolution of OST at 198 °C occurs as the sample melts.

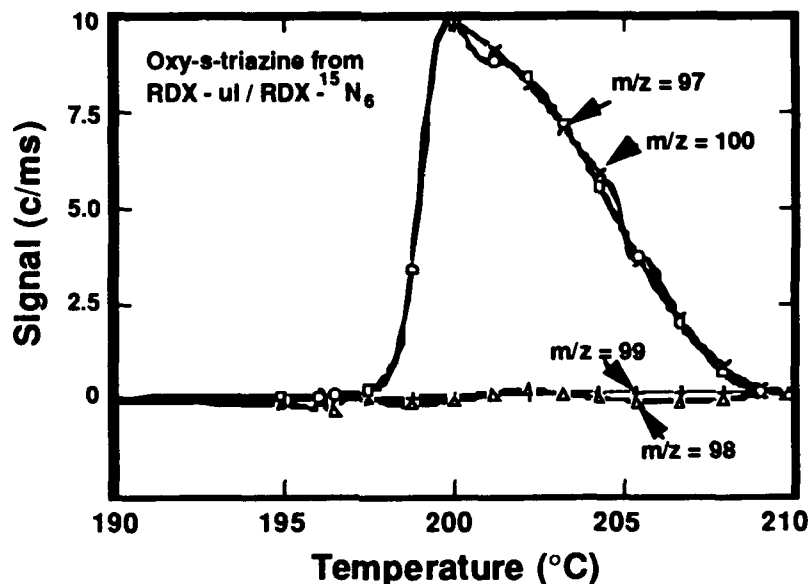


Figure 4. The ion signals from oxy-s-triazine and its isotopic analogues formed during the decomposition of a mixture of RDX-ul and RDX- $^{15}N_6$ (Experiment 3 in Table I). Between 195 and 210 °C the heating rate of the sample is 0.5 °C/min. The rapid increase in the rate of evolution of OST at 199 °C occurs as the sample melts.

analogues. However, scrambling of the ^{13}C and ^{18}O in formaldehyde may be the result of polymerization and depolymerization of formaldehyde after its initial formation and prior to its detection. The ion signals of $^{13}\text{CH}_2\text{O}$ (31), CH_2^{18}O (32), and $^{13}\text{CH}_2^{18}\text{O}$ (33) were used to calculate the value for f .

N_2O . Experiment 3 (Table IV) with a mixture of RDX-ul and RDX- $^{15}\text{N}_6$ shows that none of the N_2O is formed from mixing between the two differently labeled RDX analogues. The ion signals of $^{14}\text{N}_2\text{O}^+$, $^{14}\text{N}^{15}\text{NO}^+$, and $^{15}\text{N}_2\text{O}^+$ (Fig. 2) are used to calculate the mixing fraction.

Table IV. N_2O Isotopic Scrambling Results

Isotopes	Exp #	m/z	Rel Intensities ^a	Temporal Corr
$^{14}\text{N}/^{15}\text{N}$	3	44	100 ± 10	Y
		45	-4.1 ± 2	na
		46	93.9 ± 11	Y

^a The ion signal at each m/z value is corrected for overlapping ion signals that are observed in experiments with individual RDX-ul and RDX- $^{15}\text{N}_6$ samples.

OST. Three different kinds of isotopic scrambling experiments were used to derive mixing information for OST. The results are tabulated in Table V and shown in Figures 3 to 5. Analysis of the data presented in Table V shows *no mixing in any of these experiments*.

The experiments with RDX-ul and RDX-d₆ (Fig. 3, Table VII) show a strong DKIE. The relative size of the DKIE ranges from 2.2 during the initial stage of formation of OST down to 1.3 at the final stage of formation of OST. This decrease in the DKIE is probably due to the increase in the relative amount of RDX-d₆ compared to RDX-ul as the sample decomposes.

ONDNTA. ONDNTA represents a product (M.Wt. 206) that is formed by the net loss of a single atom (oxygen) from RDX. The results from two different kinds of isotopic scrambling experiments for ONDNTA are tabulated in Table VI and the results of the $^{14}\text{N}/^{15}\text{N}$ isotopic exchange measurements from ONDNTA are shown in Figure 6. The isotopic analysis of this product from the mixing experiments was based on its major fragment ion, m/z = 132 ($\text{C}_2\text{H}_4\text{N}_4\text{O}_3$) and its isotopic variants.

The data indicate that 25 ± 5 percent of the RDX participates in scrambling of H and D in the formation of ONDNTA, in spite of the fact that the product results from the loss of a single oxygen atom. But since there is no DKIE observed in the formation of this product, hydrogen is not involved in the rate determining step for this product.

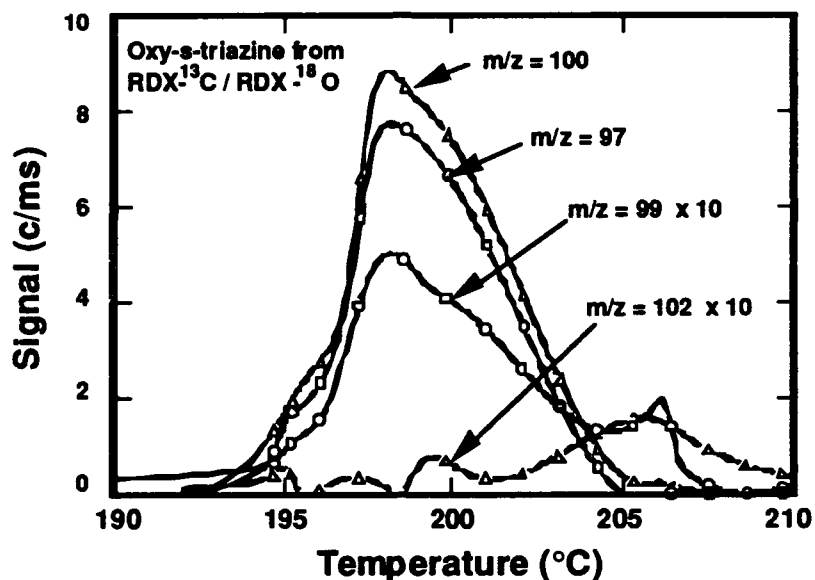


Figure 5. The ion signals from oxy-s-triazine and its isotopic analogues formed during the decomposition of a mixture of RDX- ^{13}C and RDX- ^{18}O (Experiment 4 in Table I). Between 195 and 210 $^{\circ}\text{C}$ the heating rate of the sample is 0.5 $^{\circ}\text{C}/\text{min}$. The rapid increase in the rate of evolution of OST at 196 $^{\circ}\text{C}$ occurs as the sample melts.

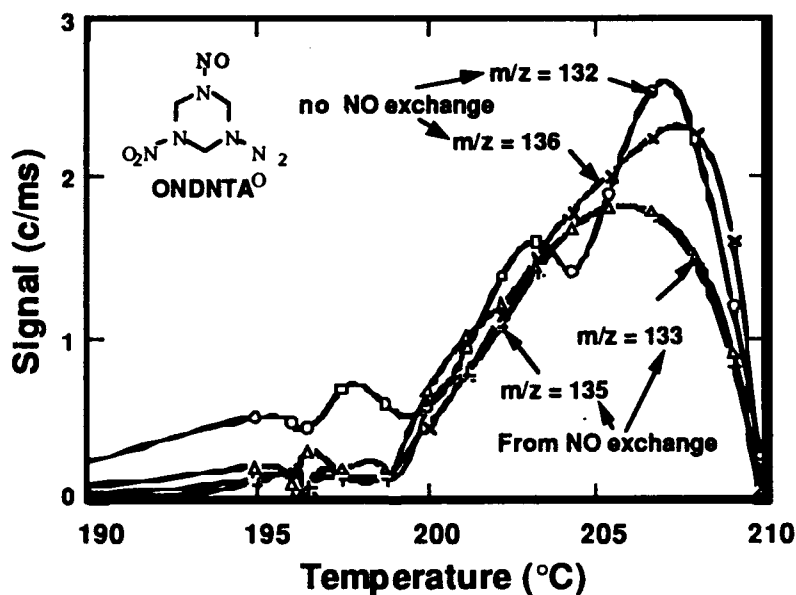


Figure 6. The ion signals, from ONDNTA and its isotopic analogues in the decomposition of a mixture of RDX-ul and RDX- $^{15}\text{N}_6$ (Exp. 3 in Table I). Heating rate of the sample is 0.5 $^{\circ}\text{C}/\text{min}$. The larger variation in the ion signal at $m/z = 132$ is due to the large correction made to the signal for the contribution of the $\text{C}_3\text{H}_4\text{N}_3(\text{NO}_2)$ ion dissociation fragment formed from evaporating RDX- ^{15}N .

Table V. Oxy-s-triazine (OST) ($C_3H_3N_3O$) Isotopic Scrambling Results.

Isotopes	Exp #	m/z	Rel Intensities ^a	Temporal Corr
H/D	1	97	100±3	N
		98	4±3	na
		99	4±3	na
		100	45 to 71±3	N
	2	97	100±2	Y
		98	3.1±2.1	na
		99	2.8±0.4	na
		100	38 to 59±2	Y
¹⁴ N/ ¹⁵ N	3	97	100±2	Y
		98	1±2	na
		99	1±2	na
		100	100±2	Y
¹³ C/ ¹⁸ O	4	97	85.4±4	Y
		99	5.4±0.5	Y
		100	100.0	Y
		102	0.8±0.8	N

^a Ion signals at m/z-values of 98 and 102 have been corrected for the ion signal due to fragmentation of RDX in the ionizer of the mass spectrometer.

The ¹⁴N/¹⁵N mixing in the ONDNTA product can be readily seen in Figure 6. During the initial stages of the decomposition, the ion signals at m/z-values of 132, 133, 135, and 136 are approximately equal, thus indicating that the mixing is complete with respect to the N-NO bond. Whereas, during the latter stages of the decomposition, the relative sizes of the ion signals at m/z-values of 132 and 136 are larger compared to the ion signals at m/z-values of 133 and 135, thus indicating that the extent of RDX that participates in mixing with respect to the N-NO bond falls to 60 percent.

Other products. Other products formed during the decomposition of RDX do not involve scrambling of bonds and are not presented here.

Discussion

As indicated in the introduction, this paper focuses attention on the deuterium kinetic isotope effects and isotopic scrambling results of the decomposition products from mixtures of isotopically labeled RDX samples (Table I). However, it is helpful and necessary to interpret these results in a broader perspective taking into account the relative amounts of all the products (see previous paper), the role of physical factors and a comparison to similar data obtained from HMX. While the use of isotopes has been helpful in identifying several product species in

these decomposition studies and the bond-making and bond-breaking processes, they also uncover some limitations in their use to draw conclusions about reaction mechanisms.

Table VI. ONDNTA ($C_3H_6N_3(NO_2)_2NO$) Isotopic Scrambling Results

Isotopes	Exp #	m/z ^a	Rel Intensities ^b	Temporal Correlation ^c
H/D	1	132	100±14	Y
		133	7±5	Y
		134	0 ^d	na
		135	5±2	Y
		136	90±5	Y
	2	132	97±19	Y
		133	7±4	Y
		134	0 ^d	na
		135	8±1	Y
		136	86±3	Y
¹⁴ N/ ¹⁵ N	3	132	90 to 98±17	N
		133	95 to 70±5	Y-135
		134	0 ^d	na
		135	86 to 69±5	Y-133
		136	85 to 94±6	N

^a The m/z values are for an ion formula of $C_2H_4N_2(NO)(NO_2)$ that is formed by the fragmentation of ONDNTA in the ionizer of the mass spectrometer.

^b The ion signals at all of the m/z values have been corrected for contributions from fragmentation of RDX in the ionizer of the mass spectrometer. In addition, the ion signals at m/z-values of 133 and 135 have been corrected for contributions from isotopomers that arise directly from the individual isotopic analogues of ONDNTA in the absence of scrambling.

^c Y indicates that the ion signals at these m/z values are temporally correlated with all the other m/z values listed in the experiment. N indicates that the ion signals are not temporally correlated with other ion signals in the group. Y-# indicates that the ion signal is correlated with another m/z-value in the group indicated by #.

^d There is no evidence of transfer of two atoms between the product molecules.

It has already been stressed in the preceding papers in this series that several decomposition channels can be identified in both HMX and RDX decompositions. In view of this it is pertinent to consider the relative importance of these channels of decomposition and understand the similarities and differences between HMX and RDX imposed by the respective structural similarities and differences in physical properties (see introduction).

Table VII. Mixing Fractions and DKIE Results from RDX Isotopic Scrambling Experiments.

Decomposition Products	Scrambled Isotopes	Mixing fraction f	DKIE x^a
H ₂ O	H/D	1.1±0.1	1.5±0.1
CH ₂ O	H/D	0.075±0.02	1.05±0.1
	¹³ C/ ¹⁸ O	0.45±0.07	
N ₂ O	¹⁵ N ₆ / ¹⁴ N ₆	-0.04±0.03	
OST	H/D	0±0.05	2.2 to 1.3 ^b
	¹⁵ N/ ¹⁴ N	0±0.05	
	¹³ C/ ¹⁸ O	0±0.05	
ONDNTA	H/D	0.25±0.05	1.05±0.2
	¹⁵ N ₆ / ¹⁴ N ₆	1.0 to 0.61 ^c	

^a The values for x are calculated only for the H/D exchange experiments.

^b The DKIE for OST decreases as the sample undergoes decomposition due to the increase in the amount of RDX-d6 relative to the amount of RDX-ul.

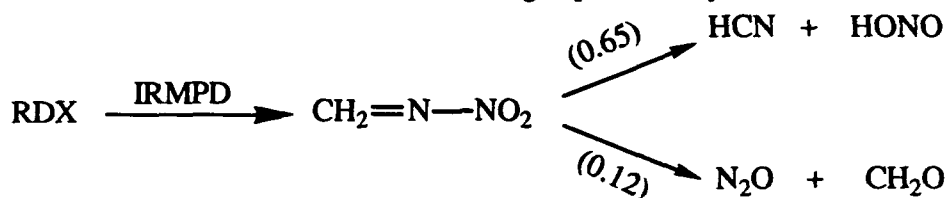
^c The fraction of NO that exchanges during the formation of ONDNTA decreases as the sample size decreases.

N₂O and CH₂O: In RDX decomposition, only a small amount of isotopic scrambling was observed for the C-H bonds of formaldehyde (Table VII and Figure 1) and the N-N bonds in N₂O (Table VII and Figure 2), which are two major products. This is consistent with previous work³ on RDX-(¹⁵NO₂)₃ in which it was found that N₂O consisted mainly of ¹⁴N-¹⁵N-O configuration. If the N-N bond is broken in the reaction path that leads to the formation of N₂O from RDX, then the results with RDX-¹⁵N₆/RDX-ul should have shown mixed isotopes. That this was not the case (Table IV) shows clearly that, as in the case of HMX, N₂O formed in the decomposition originates both directly from the parent molecule and indirectly through an intermediate (ONDNTA) without rupture of the N-N bond.

While the C-H bonds in CH₂O exhibit only a minor degree of mixing, the ¹³C/¹⁸O atoms showed 45 percent scrambling. This means that 45 percent of all the C-O bonds in the CH₂O product molecules are formed from atoms drawn from different isotopic molecules. However, this scrambling can readily occur in a CH₂O polymerization-depolymerization process any time prior to exiting from the reaction cell. In the case of HMX the degree of mixing is even higher⁵ (80%) presumably because CH₂O is held for longer periods within the solid particles compared to the decomposition of RDX in a melt. The minor amount (< 7%) of H-D scrambling, likewise could have occurred after the initial formation of CH₂O due to the labile nature of the hydrogens. The lack of a DKIE for CH₂O shows that C-H bond cleavage is not involved in the rate limiting step of the reaction pathway leading to its formation, thus providing further evidence of several parallel reaction pathways controlling the decomposition of RDX.

BEHRENS and BULUSU

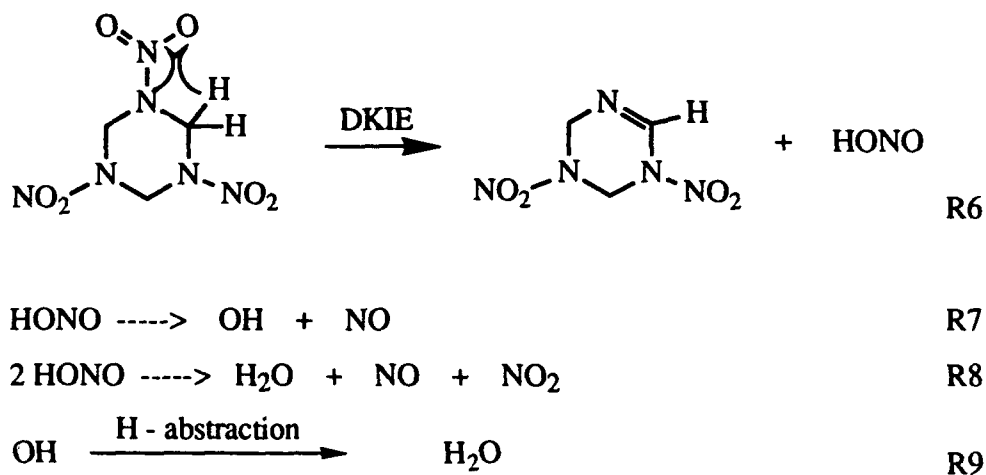
The isotopic scrambling results of N_2O and CH_2O , taken together, point in the direction of a methylenenitramine decomposition to N_2O and CH_2O (Reactions R1, R2, and R5) as in the case of IRMPD of RDX.¹⁰ However, in the IRMPD gas-phase study



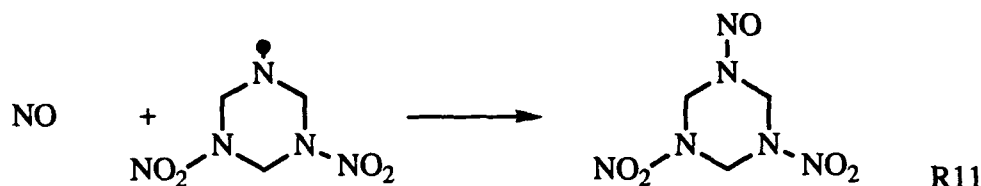
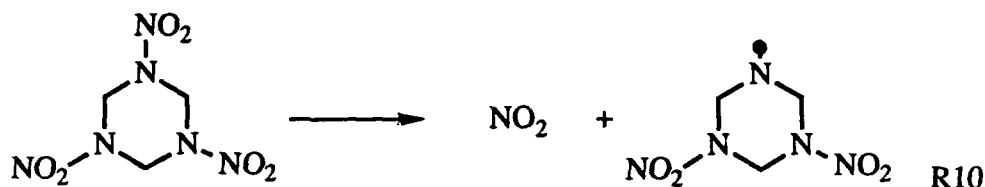
the N_2O and CH_2O channel was not as dominant. In light of the three primary decomposition pathways of RDX in the liquid phase that lead to the formation of N_2O ⁷ this difference is not surprising. Most of the N_2O formed in the decomposition of RDX in the liquid phase occurs through either the ONDNTA intermediate (~ 42%) or through the catalyzed decomposition of RDX (~ 42%). Only a small fraction (~ 15%) of the N_2O appears to be formed directly from RDX. Whether the N_2O formed via either the ONDNTA pathway or the catalytic pathway passes through a methylenenitramine intermediate is uncertain at this point.

H_2O formation. The H_2O formation rate displays a strong kinetic isotope effect (Table VII) suggesting that it is formed in a reaction pathway, such as (R6), in which hydrogen is involved in the rate limiting step.

In the following scheme, reaction R6 generates HONO accompanied by a primary kinetic isotope effect with deuterium label since a C-H bond scission is necessary. HONO, under reaction conditions, can undergo one or both of the two possible reactions, R7 and R8. In the former case, a hydroxyl radical is generated which on hydrogen abstraction (R9) yields H_2O as a product.



Alternately, the decomposition of HONO, R8, can directly yield H_2O and two oxides of nitrogen all of which have been detected among the products. The latter may be expected to



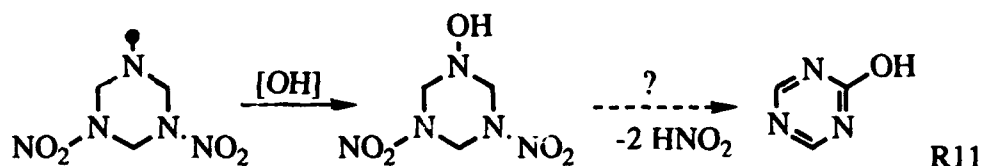
seek other radical species to react with, such as the radical left behind on loss of an NO_2 group from RDX (R10). In this case, NO_2 will reform the RDX molecule while NO will lead to the formation of ONDNTA (R10) detected in the products.

ONDNTA: It is interesting that the ONDNTA formation as a product (Tables VI and VII) shows a substantial mixing of the nitrogen isotopes making up the N-NO bond and 25 percent mixing of H/D isotopes. The former is expected from R11 but it is not immediately apparent why H/D mixing is observed. The explanation for this may be that the hydrogens next to a nitroso group become labile (acidic) and exchangeable. In fact, this exchangeable nature of the hydrogens is taken advantage of in the synthesis of the nitroso precursor to RDX- d_6 .

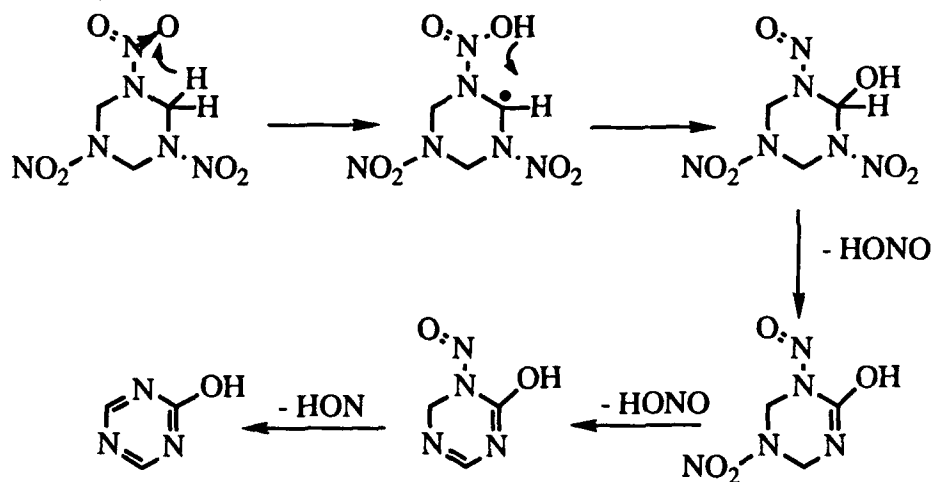
There is also NMR evidence showing a substantial contribution of the resonance structure with charge localization on the ring nitrogen atom that would make the adjacent hydrogen more acidic and susceptible to hydrogen abstraction promoting decomposition of ONDNTA in the earlier stages of the reaction⁷ due to reaction with other decomposition products.

The mixing of the nitrogen isotopes in the N-NO bond of course suggests the breaking of the N- NO_2 bond and reforming it with a nitroso group. This mechanism is quite analogous to the one proposed for the formation of dimethylnitrosamine.⁴ A radical addition, $\text{R}\cdot$, to one of the oxygens of the nitro-group followed by loss of the RO group to give nitroso-amine as proposed by Melius²⁰ is not valid here since it would not predict mixing of isotopes in the N-N bond.

Oxy-s-triazine (OST): Oxy-s-triazine has proved to be a most intriguing product from RDX to rationalize:



It would be reasonable to expect the radical produced in Reaction 10 to pick up a hydroxyl radical and eventually lead to OST by loss of 2 HNO_2 molecules and internal bond rearrangement. However, the first objection to this would be the excess hydrogen atom on the ring. But, more importantly, OST is unique in that it fails to show mixing of any of the isotopes used (H/D , $^{14}\text{N}/^{15}\text{N}$, or $^{13}\text{C}/^{18}\text{O}$). This requires that it be formed in a strictly intramolecular process. It appears, therefore, that OST represents an additional decomposition pathway that is most likely a unimolecular one. It can be seen in Figures 3 - 5 that the signal from OST abruptly increases at the melting point of RDX and then progressively decays as the sample is consumed. In order to take place, this reaction has to result in a concerted loss of 2 HONO molecules and an HNO radical by the RDX molecule. This reaction route to OST accounts for the total lack of mixing of the isotopes, the consequent intramolecular nature of the product, the unimolecular kinetics apparent from Figures 3 - 5 and the observed DKIE. It is interesting to note that OST is the only decomposition product from RDX that solely exhibits unimolecular decomposition kinetics (other products are produced by both the unimolecular and other decomposition pathways).



Summary

The results from isotope crossover experiments with RDX and the observed deuterium kinetic isotope effects (DKIE) lend further support for the four primary decomposition pathways described in a previous paper⁷ and may be summarized as shown in the scheme below. Two pathways are first order reactions solely in RDX. Pathway 1 produces predominantly OST, NO , and H_2O and accounts for approximately 30% of the decomposed RDX and Pathway 2 produces predominantly N_2O and CH_2O with smaller amounts of NO_2 , CO , and NH_2CHO and accounts for 10% of the decomposed RDX. Pathway 3 consists of formation of ONDNTA by reaction between NO and RDX, followed by the decomposition of ONDNTA to predominantly CH_2O and N_2O . Pathway 4 consists of decomposition of RDX through

	<u>Products</u>	<u>Path</u>	<u>Fraction</u>	<u>DKIE</u>	<u>$^{14}\text{N}/^{15}\text{N}$ Exchange</u>
RDX	OST + H ₂ O + NO + NO ₂	1	30%	Yes	No
	NO ₂ + H ₂ CN + 2 N ₂ O + 2 CH ₂ O	2	10%	No	No
	ONDNTA \longrightarrow N ₂ O + CH ₂ O + other	3	35%	No/Yes	Yes
	N ₂ O + CH ₂ O + NO ₂ + NH ₂ CHO	4	25%	No	No

Arrows from RDX point to each product row. The arrow to Path 3 is labeled "[NO]" and the arrow to Path 4 is labeled "Catalyst".

reaction with a catalyst that is formed from the decomposition products of previously decomposed RDX. Reaction Pathways 3 and 4 each account for approximately 30% of the decomposed RDX.

The lack of isotopic scrambling and the presence of a primary DKIE in OST support a unimolecular mechanism in Pathway 1. The primary DKIE in the formation of OST shows the importance of intramolecular hydrogen transfer in the rate limiting step that controls Pathway 1.

Reaction Pathway 3 is supported by the scrambling of the $^{14}\text{N}/^{15}\text{N}$ in the nitroso group of the ONDNTA formed during the decomposition of RDX. In addition, the lack of a DKIE in the formation of ONDNTA also supports this pathway. Furthermore, the presence of H/D scrambling is consistent with the presence of more labile hydrogen on the the neighboring carbon atom of the N-NO group in ONDNTA.

The lack of nitrogen scrambling in N₂O is consistent with the conclusion that methylene nitramine collapses to N₂O and formaldehyde without cleavage of the N-N bond. However, the lack of direct evidence of methylenenitramine makes its formation through this intermediate uncertain.

The lack of a DKIE in the formation of CH₂O supports the conclusion that hydrogen is not involved in the rate limiting steps of Pathways 2, 3 and 4 that lead to its formation.

Unlike the case of HMX decomposition which is catalyzed by H₂O, there seems to be no evidence of autocatalysis by H₂O in the RDX case. H₂O formation is not the sole contributor to the observed DKIE as in the case of HMX. It is, therefore, apparent that the decomposition details of RDX and HMX are significantly different due to the fact that the former occurs in the

liquid phase and the latter in the solid phase. Furthermore, RDX undergoes isotopic scrambling of nitrogen in the formation of ONDNTA, whereas for the analogous compound in HMX, ONTNTA, only 25% of the molecules are formed by mixing of the nitrogen isotopes. This implies that the formation of ONTNTA in HMX is influenced by the "cage effect" in its solid state decomposition, whereas the formation of ONDNTA in RDX occurs in the liquid phase and the 'cage effect' does not play a role. This also predicts that if HMX is decomposed above its melting point, it also would exhibit more scrambling of the nitrogen isotopes.

References

- (1). For reviews see a) T.L. Boggs, "The Thermal Decomposition Behavior of Cyclotrimethylene - trinitramine (RDX) and Cyclotetramethylene - tetranitramine (HMX) in *Fundamentals of Solid-Propellant Combustion*," Kuo, K.K. and M. Summerfield eds.; *Progress in Astronautics and Aeronautics* Vol. 90., p. 121, AIAA Inc., New York, NY, 1984; b) M.A. Schroeder, "Critical Analysis of Nitramine Decomposition Data: Product Distributions from HMX and RDX," BRL-TR-2659, 1985; d) M.A. Schroeder, "Critical Analysis of Nitramine Decomposition Data: Activation Energies and Frequency Factors for HMX and RDX Decomposition," BRL-TR-2673, 1985.
- (2). S. Bulusu, and R.J. Graybush, "Proceedings of the 36th International Congress on Industrial Chemistry Brussels, Belgium, 1967; *C. R. Ind. Chim. Belge*, **32**, 647 (1967).
- (3). R.J. Karpowicz, and T.B. Brill, *AIAA Journal*, **20**, 1586 (1982).
- (4). R. Behrens, Jr., *J. Phys. Chem.*, **94**, 6706, (1990).
- (5). R. Behrens, Jr. and S. Bulusu, *J. Phys. Chem.*, **95**, 5838 (1991).
- (6). R. Behrens, Jr., "Proceedings of 23rd JANNAF Combustion Meeting", Langley, VA, CPIA Publ., 457, pp. 231-240, (1986).
- (7). R. Behrens, Jr. and S. Bulusu, *J. Phys. Chem.*, (in press).
- (8). A.P. Snyder, J. H. Kremer, S.A. Liebman, M.A. Schroeder, and R.A. Fifer, *Org. Mass Spec.*, **24**, 15 (1989).
- (9). C. F. Melius, *J. de Physique, Coloque C4*, **48**, 341 (1987).
- (10). X. Zhao, E.J. Hints, Y.T. Lee, *J. Chem. Phys.*, **88**, 801 (1988).
- (11). R. Shaw, and F.E. Walker, *J. Phys. Chem.*, **81**, 2572 (1977).
- (12). a) R. Behrens, Jr., *Rev. Sci. Instrum.*, **58**, 451 (1986); b) R. Behrens, Jr., "The Application of Simultaneous Thermogravimetric Modulated Beam Mass Spectrometry and Time-of-Flight Velocity Spectra Measurements to the Study of the Pyrolysis of Energetic Materials," in *Chemistry and Physics of Energetic Materials*, Bulusu, S. N., Ed.; *Proceedings of the NATO Advanced Study Institute*, Vol. 309, Kluwer Academic Publishers, Netherlands, 1990, p. 327.
- (13). R. Behrens, Jr., *Int. J. Chem. Kinetics*, **22**, 135 (1990).
- (14). S. Bulusu, J.R. Autera, and T. Axenrod, *J. Labeled Compounds & Radiopharma.*, **17**, 707 (1980).
- (15). J.H. Boyer, and G. Kumar, *J. Labeled Compds. & Radiopharma.*, **22**, 1 (1985).
- (16). F.J. Brockman, D. C. Downing, and G.F. Wright, *Can. J. Research*, **27B**, 469 (1949).
- (17). R. Behrens, Jr., *Int. J. Chem. Kinetics*, **22**, 159 (1990).
- (18). S. Bulusu, J.R. Autera, and R.J. Graybush, "Proceedings of the 1968 Army Science Conference (OCD), West Point, N.Y., **2**, 423 (1968).
- (19). a) S.E. Nigenda, D.F. McMillen, and D.M. Golden, *J. Phys. Chem.*, **93**, 1124 (1989). b) J.M. Fluornoy, *J. Chem. Phys.*, **36**, 1106 (1962).
- (20). C.F. Melius, and J.S. Binkley, In *Proceedings of the 21st Symposium (International) on Combustion*, 1987; The Combustion Institute, Pittsburgh, PA, 1988, p. 1953.

BERGER

A New Understanding of Power Transmission
and Reflection in Remote and Near-In Sensing

Dr. Henry Berger*
Army Space Technology and Research Office
U.S. Army Topographic Engineering Center
ATTN: ASTRO (SLCSP)
Ft. Belvoir, Virginia 22060-5546

I. INTRODUCTION.

This paper proposes a new understanding of power transmission and reflection for remote and near-in sensing through, and of, inhomogeneous media (IM). The formulation developed is also directly applicable to single mode Nonuniform Waveguiding Structures (NOWS) and extendable to multi-mode NOWS. The discussion focuses on how these processes are characterized by the classical generalized transmission and reflection coefficients, $T(x)$ and $R(x)$, respectively, where x denotes a spatial variable. It was prompted by the very recent discovery, by this author, that two of the classical formulations for $R(x)$, in widespread usage by the international electromagnetics and acoustics communities for decades, appear to be flawed.¹ These formulations are currently used throughout a broad spectrum of technology, engineering and scientific applications, briefly discussed below.

Since the two formulations in question have been in widespread use for so long, a key question is: If they are truly flawed wouldn't that have been detected long ago? Surprisingly, there are a number of good reasons why they were not and the discussion below includes an additional very important one, not noted in reference 1. A companion question is: Can an experiment be devised that can accurately check the validity of the competing formulations in a way that will satisfy the proponents of each? This author proposes below a technique he believes will do that.

However, there is something more to be tested because the new formulation suggests a new physical paradigm for understanding where and how the reflection and transmission processes take place within the IM. And the proposed experiment appears to also be capable of verifying whether this new interpretation is soundly based.

Inhomogeneous media are present in most regions where there is an interest in performing remote sensing, such as in low earth orbits (LEOs). LEO is variously defined, for near circular orbits, to be orbits between the altitudes of 150 km and 900 km, while the ionosphere is variously described as extending from an altitude of 50 km up to 1000 km. Thus, LEOs generally lie somewhere within the ionosphere, a highly inhomogeneous medium.

The earth is a medium, with inhomogeneities in a variety of constituents near and at the surface, challenging earth penetrating radar trying to probe beneath that surface. Surface layers of vegetative canopies, ice, snow, mud, water, and sand present different levels of inhomogeneity to probing signals from up close and from afar. Überall and his colleagues explained over a decade ago how to determine material composition of penetrable homogeneous radar targets of a given simple shape from afar and described a method for determining layer thickness, density, sound speed and absorption coefficient of homogeneous viscous fluid layers with acoustic echos.^{2,3} Extending those accomplishments to inhomogeneous layers will probably require another approach but when that is accomplished it will be of great value.

Generally, it is understood that in remote sensing the sensor system cannot control the near-in environment surrounding the object of interest. One might then define near-in sensing as being performed on objects near enough to control their immediate environments totally or to a significant degree. One example is nondestructive probing of solid-state devices for their parameter values in key regions. This can not only be of great value during the research and development phase of programs, but in quality control during the production phase. Inverse scattering techniques used for probing the ionosphere have been applied to determining spatial variability of the index of refraction within optical waveguides. An additional example of near-in sensing of inhomogeneous media is medical imaging of human tissue using diffractive tomographic imaging.

Another cluster of applications involves the synthesis, analysis and design of nonuniform, distributed electromagnetic circuits and circuit components. Womack demonstrated over three decades ago a length reduction of 40% in coaxial resonant cavities, by using exponential nonuniformities, without noticeable degradation in performance.⁴ This differs

from the size reduction in microwave integrated circuits (MIC) and monolithic MIC which involve reduction in transverse dimensions, not length. Because of this, tapered transmission lines continue to be used and show up in the latest design literature.⁵

II. BACKGROUND.

Reflection coefficient, as it is commonly understood for homogeneous media (HM) and uniform waveguiding structures (UWS) measures power reflected (P_r) from some interface, expressed by

$$|R|^2 = P_r/P_i, \quad (1)$$

where P_i is the incident power, $|R|$ represents the absolute value of R and the fields are time-harmonic. The term generalized reflection coefficient is sometimes used to emphasize that R is a function of position, x .

In practical applications equation (1) is probably the most widely used property of $R(x)$. For HM and UWS,

$$R(x) = [Z(x) - Z_c] / [Z(x) + Z_c], \quad (2a)$$

where $Z(x)$ is the impedance at x and Z_c is the characteristic impedance of the medium or waveguide and is constant. The inverse relationship is,

$$Z(x) = Z_c [1 + R(x)] / [1 - R(x)]. \quad (2-b)$$

Schelkunoff appears to have been the first to have discussed, in general and at length, the concept of reflection coefficient for IM and Nows.⁶ His formulation deals only with the value of the reflection coefficient, $R_s(x_b)$, on the boundary, at $x = x_b$, separating some homogeneous medium from an inhomogeneous medium. Schelkunoff asserted that,

$$R_s(x_b) = [Z(x_b) - Z_c(x_b)] / [Z(x_b) + Z_c(x_b)], \quad (3)$$

where $Z(x_b)$ is the nominal characteristic impedance of the medium at $x = x_b$. Later, Walker and Wax, in 1946, asserted the more general relationship in equation (4) held,⁷

$$R_w(x) = [Z(x) - Z_c(x)] / [Z(x) + Z_c(x)] \quad (4)$$

They offered no argument as to why equation (4) was valid, they cited no empirical evidence, and they referenced no prior literature on the subject. Nevertheless $R_v(x)$ has become widely accepted and used in both the electromagnetics and acoustics communities in the four decades since then.

For both $R_v(x)$ and $R_s(x_b)$ it has always been understood, without proof, that equation (1) still holds and it is so used in the literature. There is a third formulation of reflection coefficient by Epstein (R_e) that applies to problems with diffuse boundaries such as the top of the atmosphere.⁸ R_e has been shown to be consistent with power conservation. Because all these formulations are generally taken to be describing different facets of the same entity, the literature does not discriminate between them, neither by subscript nor name. It has been simply assumed that they are all consistent with power conservation.

In 1950, in an apparently little known conference paper, Schelkunoff argued that $R_v(x)$ was "...meaningless..." except on the boundary, x_b .⁹ Schelkunoff's argument differs from that in reference 1. R_e is not necessarily correct even for lossless IM/NOWS which have continuity of nominal characteristic impedance, $Z_c(x)$, due to problems with boundary data transformations as discussed in reference 1. For IM and NOWS with discontinuities in $Z_c(x)$, R_s is always incorrect, though in many cases the error can be small. This will be discussed in detail in a future publication.

Schelkunoff appears to argue that the forms of the elementary travelling waves in HM and UWs waves are crucial to the identification of reflected and transmitted power (P_t). Here,

$$E_y(x) = A \cdot \exp[-jkx] + B \cdot \exp[+jkx], \text{ and} \quad (5)$$

$$H_z(x) = Y_c \cdot A \cdot \exp[-jkx] - Y_c \cdot B \cdot \exp[+jkx], \quad (6)$$

where $Y_c \equiv 1/Z_c$, $Z_c = \sqrt{\mu/\epsilon}$, A and B are amplitude constants, $k = \omega/\mu\epsilon^c = \text{constant}$, $\omega = 2\pi f = \text{radian frequency}$, μ and ϵ are the permeability and permittivity, respectively, of the medium, E_y is the y-component of the electric field, H_z is the z-component of the magnetic field and $\exp(y) \equiv e^y$.

For simple TEM, time-harmonic waves, the transmitted

power per unit area, in the x-direction is given by,

$$P_t = \frac{1}{2} R_e \{ E_y(x) \cdot H_z(x)^* \} = P_i [1 - |R(x)|^2] = \text{constant}, \quad (7)$$

where P_i = incident (forward-travelling) power = $\frac{1}{2} Y_c |A|^2$ and the reflection coefficient is given by,

$$R(x) \equiv R_0 \exp[+2jkx] \quad , \quad R_0 \equiv B/A. \quad (8)$$

While this author agrees that R_0 is meaningless, he does not agree that the simple forms in equations (5) and (6) are necessary to identify the reflected and transmitted components of power.

III. IDENTIFYING COMPONENTS OF POWER.

In reference (1) this author considered the simplest IM/NOWS case where the phenomena of interest occur. That is the scalar problem of a time-harmonic, electromagnetic, TEM mode propagating in lossless IM, or lossless single-mode NOWS characterized by,

$$Z_c(x) = Z_{co} \cdot \exp[2gx], \quad 0 < x < L, \text{ and} \quad (9)$$

$$k = 2\pi/\lambda = 2\pi f/c \quad , \quad Y_{co} \equiv 1/Z_{co}, \quad (10)$$

where g = constant taper rate and $Z_{co} = Z_c(0) = \text{constant}$.

In field notation, the results were,

$$E_y(x) = A \cdot \exp(gx) \cdot \exp(-jpx) [1 + R_p(x)], \quad (11)$$

$$H_z(x) = Y_{co} \cdot A \cdot \exp(-gx) \cdot \exp(-jpx) [a - a^* R_p], \quad (12)$$

$$R_p(x) \equiv R_{po} \cdot \exp[+2jpx], \quad R_{po} \equiv B/A \quad (13)$$

$$p^2 \equiv K^2 - g^2, \quad \bar{g} \equiv g/k, \quad \bar{p}^2 \equiv p^2/K^2 = 1 - \bar{g}^2, \text{ and} \quad (14-a)$$

$$a \equiv \bar{p} + j\bar{g}, \quad a^* \equiv \bar{p} - j\bar{g}. \quad (14-b)$$

The impedance was given by,

$$Z(x) \equiv E_y/H_z = Z_{co} \cdot [1 + R_p(x)]/[a - a^* R_p(x)]. \quad (15)$$

Elaborating on the calculation of transmitted power which yields,

$$P_t = \frac{1}{2} Y_{co} \cdot \bar{p} |A|^2 \{1 - |R_p(x)|^2\}, \quad (16)$$

it is clear that equation (16) can be rewritten as,

$$P_t = P_f - P_r, \quad (17)$$

where the forward and reflected travelling components of the power can be identified as,

$$P_f = \frac{1}{2} Y_{co} \bar{p} |A|^2, \text{ and} \quad (18)$$

$$P_r = R_p(x) \cdot P_f, \quad (19)$$

respectively. It is clear that P_f is independent of the load parameters and that it reduces smoothly to P_i = incident power in equation (7) as \bar{g} (the normalized constant taper rate) goes to zero. On the other hand, from equations (16), (17) and (19), it can be shown that,

$$R_p(x) = R_l \cdot \exp[+2jp(x - L)], \quad (20)$$

where R_l is the load reflection seen from the inside edge of the region $0 < x < L$ which will be denoted by L^- . The choice of $x = L^-$ as opposed to $x = L$ is forced by the occurrence of a discontinuous change in reflection coefficient as the boundary is crossed (see reference (1)). It also can be shown that,

$$0 \leq |R_p(x)| \leq 1. \quad (21)$$

Thus, if $R_l = 0$, then from equations (16) and (17) it can be seen that $P_t = P_f$, ie. all the available power in P_f is transmitted. When $|R_l| = 1$, then equations (16) and (17) reveal that $P_t = 0$, no power is transmitted. As $|R_l|$ is swept through the numerical range from zero to one, P_t is swept through the range from its maximum possible value, P_f , to its minimum possible value, zero.

The description in the preceding paragraph, of how P_t varies as R_l and hence $|R_p|$ varies is identical to that of how it varies for HM and UWS. Thus R_p has the identical properties, within this framework, as the reflection coefficient defined for HM and UWS.

There is an additional similarity between $R_p(x)$ in exponential IM and $R(x)$, one that is intuitively satisfying,

but not to be expected in most complex IM and NOWS. It is that, exactly as for $R(x)$, $R_p(x)$ can be shown to be,

$$R_p(x) = \frac{B \cdot \exp[j(\omega t + px)]}{A \cdot \exp[j(\omega t - px)]} = \frac{\text{forward-travelling wave}}{\text{backward-travelling wave}}. \quad (22)$$

Now, it is reasonable to compare equation (7) and the inverse relationship,

$$R_p(x) = [aZ(x) - Z_c(x)] / [a^*/Z(x) + Z_c(x)], \quad (23)$$

with those for HM and UWS which are given in equations (2-a) and (2-b). The relationships are very similar but not identical. However as the normalized constant taper rate, g , goes to zero, a and a^* go smoothly to 1, and the two sets of relationships become identical. The above seems all that should be reasonably required to identify a reflection coefficient and reflected power. From equations (7) and (16) it appears that thinking of reflected power as "subtractive" power (so long as the forward - traveling power component, P_f , is independent of load conditions) can be used in identifying reflection coefficients in more complex problems, and this author has done so in preliminary calculations and obtained, qualitatively, the same type of results.

IV. POWER CONSERVATION AND R_w .

It is clear from comparing $R_p(x)$ and $R_w(x)$, equations (4) and (21), that the Walker-Wax formulation differs from the power formulation, for exponential variations. And although it can be shown that,

$$1 \leq |R_w(x)| \leq 1, \quad (24)$$

R_w does not appear to possess any further similarities with the $R(x)$ used in the HM/UMS case. In reference (1) it was shown that $R_w(x)$, for the exponential case, did not satisfy the weak form of power conservation (which was derived for a reflection coefficient for any spatial variation of $Z_c(x)$ and $k(x)$). The requirement of the weak form is,

$$|R(x)| = \text{constant}, \quad (25)$$

which is based, in part, on the assumption that P_f is constant because it is the power supplied by whatever source is

energizing the system and it is constant. I recognize that many advocates for the validity of R_v conceptualize the reflection and transmission process as an infinite number of multiple reflections originating from an infinite number of differential changes in $Z_c(x)$ all lined up in tandem, and so would not agree that P_r need to be constant.

But this author described in reference (1) an alternative physical picture suggested by the mathematics where, as in the HM/UWS case, only the terminal impedance at the load end of the medium or structure, produced a reflection. And a "matched" load is defined in this new context as one which does not produce reflected power, i.e. $R_p = 0$.

To apply the requirement in eg. (23) to any arbitrary spatial variations (but physically possible) in $Z_c(x)$ and $k(x)$ the standard Riccati equation for R_v can be used. Simple analysis of this yields an $R_v(x)$ which has a constant phase as well as the constant amplitude. This is an unphysical result for a medium or structure of non-zero length.

V. WHY NOT ALREADY DETECTED EXPERIMENTALLY?

It would not be correct to assume that some kinds of discrepancies, between theory and experiment have not already been published. They have. For example, in a 1990 paper by Tsedilina, in which he compares analysis with test results for a specific problem, he states "...These losses (~ 15 to 20 dB) are not described by the theory of propagation of radio waves in a smoothly inhomogeneous ionosphere..."¹⁰. While a discrepancy, for which the ratio of measurement to prediction is about 100 (a 10,000% error) is not to be regarded lightly, researchers usually do not publish negative results without some reasonable explanation, and Tsedilina is no exception. He offers a qualitative argument for some reasonable physics that he believes should be in the conventional analysis but usually isn't and his explanation may even account for most or all, of that error.

This paper does not claim that the new point of view will necessarily explain all or even any of the major discrepancies that are in the literature, only that there are some questions to be answered.

One would expect that, since measurements carried out in

the controlled environment of the laboratory tend to have the greatest accuracy, it is there that any validations be sought. However, as mentioned in reference (1), there are some difficulties. One is that the use of $R_w(x)$, even if it were correct, involves errors from assuming the presence of a single mode. The errors found experimentally for R_s might be attributed to the single-mode approximation and not the formulation of R_w itself. Another is that there are parameter ranges for which $|R_p| \approx |R_w|$.

Two such parameter ranges mentioned in reference (1) were $|R_w| \approx 1$ (as in designing resonant cavities), and, usually, $|R_p| \ll 1$ (as in designing transitions and impedance transformer sections). These represent a large fraction of reflection coefficient usage in the design of Nows. However, there is another way to look at this situation that was not discussed before. It is to ask where is the sensor and where is the emitter. A monostatic radar is an example of an active co-located sensor and emitter. When both are located within IM/Nows, or when one is and the other is not, it is straight forward to show for the lossless case, from power conservation, that there can be significant errors even in the case where only the grossest type of information is sought, such as how much power is being radiated by the emitter. First, the conventional solution, using the transmission coefficient which can be obtained from $R_w(x)$ via,

$$|T_w(x)|^2 = 1 - |R_w(x)|^2, \quad (26)$$

can be derived in the two-region case from power conservation using,

$$|T(0)|^2 P_{mw} = P_{t1} = P_{t2} = |T_w(0)|^2 P_s. \quad (27)$$

Here P_s is the power supplied by the emitter in region 2 (assumed IH/Nows), P_{mw} is the power, calculated using T_w , delivered to the sensor in region 1 (assumed HM/UWs), $x = 0$ is at the boundary between regions 1 and 2, and $T(x)$ is the transmission coefficient in the HM/UWs. If the power formulation were used instead, then equation (25) would read,

$$|T(0)|^2 P_{mp} = P_{t1} = P_{t2} = |T_p(0)|^2 P_s, \quad (28)$$

where

$$|T_p(x)|^2 = 1 - |R_p(x)|^2, \quad (29)$$

and P_{mp} is the power, calculated using T_p , delivered to the sensor. From equations (25) and (26) the normalized (fractional) error would be,

$$(P_{mp} - P_{mw}) / P_{mp} = 1 - |T_w(0)|^2 / |T_p(0)|^2. \quad (30)$$

By selecting parameter ranges where $|T_w(0)|^2 / |T_p(0)|^2$ are significantly different, experimental tests should allow a determination of which formulation is correct at the boundary, $x = 0$. However, at the boundary $T_w(0) = T_s(0)$.

Section VII will describe an experimental procedure which will allow $T_w(x)$ and $T_p(x)$, to be examined for within IM. The last portion of this section examines the three region case, a very important case for three reasons. First, it is very common both in remote sensing and in waveguide circuits. Second, it is a situation in which it is possible that the use of $T_w(x)$ instead of $T_p(x)$ will not produce any error. Third, because it is so common, it is probably the primary reason that the deficiencies of the Walker-Wax and Schelkunoff formulations have not previously come to light through experimental activity.

In the three region case there are two basic configurations of considerable interest. In both, the source and the observer lie in HM or UWS, while the propagation path between them passes thru a region of IM or NOWS. An example of when the emitter and sensor are on opposite sides is a ground based sensor from whose output signals information is sought about some characteristics of a satellite-based emitter above the ionosphere. An example of when sensor and emitter are on the same side is a ground-based passive ranging/locating station receiving a high frequency (HF) communication signal from a distant ground-based emitter when the signal is partly reflected off the ionosphere. A very common example of the three-region case for waveguides is where a tapered waveguide is used to connect a smaller UWS with a larger one.

Assume the emitter is in HM in region 1, region 2 is IM, and the sensor is in HM in region 3. Assume that region 2 is lossless and well matched to regions 1 and 3 so that it produces negligible reflections in them. Finally assume that

the boundary between regions 1 and 2 are at $x = 0$, and the boundary between regions 2 and 3 are at $x = L$. Then, from Poynting theorem we have,

$$\text{i.e.,} \quad P_{t1}(0) = P_{t2}(0) = P_{t2}(L) = P_{t3}(L), \quad (31)$$

$$P_{t1}(0) = P_{t3}(L). \quad (32)$$

When the media in region 2 is lossy in a frequency dependent manner, measurements over many frequency bands, as with multi/hyper/ultra spectral sensors, may be impacted significantly. However, when power and power-related aspects of the signals are not the primary focus, then the lossless IM can have significant impact, such as producing the equivalent of some types of multipath effects.

VI. SEPARATING BACKWARD AND FORWARD TRAVELLING WAVES.

In this section an experimental technique is proposed which appears to be able to sort out backward travelling waves from forward-travelling waves within IM as well as on their boundaries. Since what constitutes the forward and backward travelling waves in IM is at issue, the technique is designed to settle that question in a manner that is independent of which point of view is correct.

In this approach, a two-conductor, uniform hollow cylindrical waveguide of arbitrary cross-section is filled with a medium whose electrical parameters are tapered in a precise way along the cylindrical axis of the waveguide. An example of the waveguide would be the familiar uniform coaxial line with circular cross section. The two-conductor waveguide was chosen because a TEM mode may be used, as well as TE and TM modes.

Suppose that such a waveguide was filled with an inhomogeneous medium in such a way as to produce the exponential varying IM discussed earlier. There is a relatively straightforward way to measure the forward-traveling wave entering the line at one end of the waveguide, say at $x = 0$, as well as the backward traveling wave exiting at $x = L$. It is assumed that the line is terminated at $x = L$ with some desired load impedance.

A three-port circulator, described below, can be used to

perform the above measurement in a familiar way. Assume the circulator ports are numbered so that when a signal is injected into it at port one, it exits at port two. A signal injected at port two exits at port three. This is a nonreciprocal three-port device which is (ideally) matched at every port so that no part of an injected signal is reflected.

Assume a signal source with a calibrated output is connected to port one, the exponential line to port two and a particular directional coupler (terminated so that signal amplitude and phase can be measured) at port three. Injecting a signal from the signal source, and measuring the signal exiting at port three completes the task.

The challenging task is to accomplish the same sorting of backward and forward waves at some axial point, x , where $0 < x < L$. To do this a waveguide as above is needed, except that it must be fabricated such that it can be disassembled into two sections at a chosen point, say $x = b$. Call the section from $x = 0$ to $x = b$, S-I and that from $x = b$ to $x = L$, S-II. Then S-I is attached, at $x = 0$, to the three port circulator at port two, with the same attachments as before at ports one and three.

The second connection of S-I at $x = b$, is to a four-port circulator. Let the ports of the four-port be designated in the following way. A signal injected at ports a, b, c, or d exits at ports b, c, d, and a, respectively. Attach S-I at $x = b$ to port a. Attach to b a hollow uniform waveguide (W-1) to which a, say, 30 dB directional coupler is attached. W-1 is terminated in a lossless moveable shorting plate. Attach S-II, at $x = b$, to port c, and a waveguide arrangement, W-2, identical to that with W-1, to port d.

The operation is in two stages. In stage I, reference values are recorded. The two sections of waveguide are bolted together to form a single continuous line from $x = 0$ to $x = L$, and it is attached at $x = 0$ to port two of the three port circulator. A reflectionless device to measure signal output is attached to the waveguide at $x = L$. When a selected signal is injected at port one the outputs (amplitude and phase) from port three and from the waveguide at $x = L$ are measured and recorded.

In stage two the intent is to measure the forward and

backward travelling waves in the inhomogeneously filled waveguide at $x = b$, by removing them from it and injecting them into uniform waveguides where it is clear which wave, forward or backward, is being measured. They are then routed back into the inhomogeneously-filled waveguide.

The purpose of the moveable, lossless shorting plates is that when they are suitably positioned, the traveling waves (with unchanged amplitude since the hollow waveguide and shorting plates are designed for negligible attenuation) will be injected back into the inhomogeneously-filled waveguide at the same amplitude and phase they would have possessed when S-I and S-II were joined at $x = b$ into a seamless whole.

The purpose of measuring amplitude and phase of the signal exiting the inhomogeneously-filled waveguide at $x = L$ and that exiting from port three of the three-port waveguide circulator is to insure that S-I and S-II when operated with the four-port circulator have been brought back to the same operating point as when they were joined.

VII. CONCLUDING REMARKS

Currently the reflection coefficient concept can be found in the acoustic as well as the electromagnetic literatures. In electromagnetics it is used across the ultraviolet (UV), visible, infrared (IR), millimeter wave (mmw), microwave, ultra high frequency (UHF) bands and even below. Where it is used for IM or NOWS, as long as it is the one-dimensional case, the comments in this paper are applicable. An important aspect of the problem is how to identify reflection coefficients for other complex media. There is no simple answer for this, but the author is in the process of completing analysis of signal propagation in some other media types where the same kinds of results appear accessible by using the notion of subtractive power. These will be discussed in future publications. However, no general technique has been found which always produces a reflection coefficient that leads to a subtractive power description.

BERGER

REFERENCES

1. *Image Understanding for Aerospace Applications* H. Berger and D. Del Bosque, SPIE, Vol 1521, (1991), pp. 117-130.
2. *Nature*, G. C. Gaunard and H. Überall, 287, 708, (1980).
3. *J. Acoust. Soc. Amer.*, R. Fiorito, W. Madigosky and H. Überall, 1981.
4. *IEEE Trans. on Micr. Theo. Tech.*, C. P. Womack, 10, 124, 1962.
5. *Micro. J.*, J. M. Schellenberg and H. Yamasaki, 25, 51, March 1982.
6. *Electromagnetic Waves*, S. A. Schelkunoff, (Van Nostrand, Princeton, 1943).
7. *J. Appl. Phys.*, L. R. Walker and N. Wax, 17, 1043, (1946).
8. *Proc. Nat'l Acad. of Sci., U.S.*, P. S. Epstein, 16, 627 (1930).
9. *The Theory of Electromagnetic Waves*, S. A. Schelkunoff, M. Kline, ed., (Dover, New York, 1965).
10. *Geomag. Aeron.*, E. E. Tsedilina, 30, 285, 1990.

Small Scale Experiments and Modeling of Laser Ignition
of Gun Propellants

Dr. Richard A. Beyer* and Mr. Arthur Cohen
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground, Maryland 21005-5066

INTRODUCTION

The objective of the Laser Ignition in Guns, Howitzers and Tanks (LIGHT) program is to determine the feasibility of replacing or augmenting liquid and solid propellant igniter systems with lasers. Laser ignition of condensed propellants depends on the optical properties of the propellants (absorption and reflection coefficients) and the beam characteristics (wavelength, flux, pulse width and shape). To accomplish the LIGHT objective it is necessary to identify the beam characteristics for a particular propellant (or energetic material) which produce the required ignition times. It would also be helpful to have a predictive model for the laser-propellant ignition process. This paper presents the initial results of two ongoing investigations of laser ignition of solid and liquid propellants. Most of the experiments were done in closed bombs initially at atmospheric conditions. The first part of this study to be discussed is primarily concerned with evaluating ignition models. A low power carbon dioxide laser was used and loading densities were low ($\sim .0001$ g/cc) leading to relatively long reaction times. The effects of the laser beam power and pulse length and ambient gas composition were investigated. Correlations of measurements with a condensed phase ignition model are discussed. The second part of this study is concerned with identifying the laser beam characteristics appropriate for gun ignition systems. A high power neodymium laser was used and loading densities were $\sim .1$ g/cc (similar to high pressure bombs) leading to short ignition delays. The pressure records during ignition of two solid and one liquid propellant will be presented.

I. EVALUATION OF IGNITION MODELS

Radiant ignition models¹ have previously been developed for solid propellants. It is generally difficult to evaluate the usefulness of ignition models for two main reasons. The first problem is in relating ignition criteria to experimental observations. The second difficulty is that for any particular propellant neither the values of physical properties nor heat release rates which are required for the calculations are well known. Laser ignition models may have an additional problem, i.e., describing the details of the laser energy interactions (thermal and non-thermal) with the propellant (and with pyrolysis products). The adiabatic "approximate" thermal radiative ignition model described in Reference 1 has been adapted to compare with experimental data. The model considers condensed phase reactions and in-depth absorption but neglects energy losses (i.e., those losses due to surface heat transfer and emission), propellant consumption, and the effect of laser interruption (deradiation). The inability of this model to predict behavior of many solid propellants which are of interest to the LIGHT program (M30, JA2 and LOVA type) suggests that additional complexities must be included in the further development of a more complete model.

Experimental

A continuous wave CO₂ (10.6 μm) laser (Synrad model #57) with variable power (nominal 10 to 100 watts) was used as the ignition source for the solids. The response of several propellants at different power (flux) levels and in different gaseous environments has been briefly examined. Data have been obtained for the following propellants: M9 (homogeneous double-base), JA2 (homogeneous triple-base), M30 (composite triple-base), M39 (RDX composite) and HMX2 (experimental HMX composite). The samples were small disks approximately 3/16" wide. The M30 and M39 disks were cut from grains (1/4" diameter, perforated, with graphite surfaces) and the others were cut from 0.25 inch diameter strands. The beam, diameter about 3 mm, was directed along the propellant axis and covered approximately 25 percent of the sample surface. Pulse lengths were varied from 0.1 to 15 s. Measurements of light emission and pressure were made using a photomultiplier (300-800 nm response) and a piezoelectric transducer, respectively. Measurements were made in open air and in a closed one liter bomb. The vessel has optical access from four sides with one containing a ZnSe window for transmitting the 10.6 μm radiation. The bomb was initially filled with air, nitrogen, or argon. Most of the experiments were done at an initial pressure of one atmosphere. An IR microscope-spectrometer (Spectra-Tech Plan 40/Matson Polaris) was used to measure propellant reflectivities and absorption coefficients for JA2, and M9 at 10.6 μm . Propellant reaction products were analyzed by gas chromatography and FTIR spectrometry.

BEYER AND COHEN

Ignition Model

The "approximate" theory for thermal radiative ignition of solids¹ has been used to predict the dependence of propellant "ignition" delays on radiative flux. Reasonable success was obtained for single and double base propellants using experimentally determined data for the thermophysical properties and heat release rates.¹ The relevant energy equation and the initial and boundary conditions used for (transparent) solids are:

$$c \frac{\partial T}{\partial t} = ac \frac{\partial^2 T}{\partial x^2} + \frac{\partial q}{\rho \partial x} + Q_z \exp(-E/RT)$$

$$q = q_0(1-f)\exp(-nx) \quad T(x,0) = T_0 \quad \frac{\partial T}{\partial x}(0,t) = 0 \quad T(\infty,t) = T_0$$

where T = temperature, T_0 = ambient temperature, q_0 = radiant flux, n = absorption coefficient, f = reflectivity, a = thermal diffusivity, ρ = density, and c = heat capacity. The last term is the Arrhenius global heat release rate during ignition (E = activation energy, Q_z = zeroth order pre-exponential factor).

The model assumes a condensed phase surface reaction. The ignition delay is calculated as the sum of (inert) "heat up" and (chemical) induction times. The "heat up" time and the corresponding surface temperature are calculated by matching (at the surface) temperatures and time derivatives obtained from analytic solutions of the energy equations corresponding to inert heating and to adiabatic reaction. The chemical induction time is obtained from the adiabatic reaction solution. Approximations to these analytic solutions (given in Reference 1) are used in this report to calculate ignition delays as a function of radiant flux. They are valid under conditions where activation energies are much greater than thermal energies and characteristic lengths for absorption are much less than those for heat conduction during heat up. In all calculations $T_0(K)=298$, and for all propellants the following values have been used for the thermophysical properties; $a(\text{cm}^2/\text{s})=1\text{E}(-3)$, $c(\text{cal/g-K})=.35$, $\rho(\text{g/cm}^3)=1.6$.

BEYER AND COHEN

Experimental Results

A. Optical Properties

The results of the initial reflectivity measurements indicate that at $10.6 \mu\text{m}$ all propellants reflect less than 10% of the incident energy. For the calculations a reflectivity value of 0.08 has been used. At present, only the M9 and JA2 absorption coefficients (α) have been determined and are equal to 250 cm^{-1} and 536 cm^{-1} , respectively.

B. Laser Ignition

The laser characteristics corresponding to the emission and pressure traces shown in the following figures and various delay measurements obtained from these traces are listed in Table I. P is the laser power. The time (relative to start of the laser irradiation) at deradiation (laser off) is denoted by t_i ; t_e is the time for start of emission or emission delay. The delay to flamespreading, which is determined from a slope change in the emission-time signal, is denoted by t_f . The exact significance of t_f needs to be determined. t_p is the time for the start of pressurization and t_d is the final emission decay time. In the figures e, f and d correspond to t_e , t_f and t_d , respectively.

Table I. Laser Operating Conditions and Delay Measurements for Figures 1, 5, 7 and 9

Figure	Signal	P (w)	t_i (s)	t_e (s)	t_p (s)	t_f (s)	t_d (s)
1	1	2.5	5.80	.19	--	.30	1.65
	2	8.6	1.10	.03	--	.09	1.40
2	1	3.2	1.08	.12	.18	.20	2.18
	2	3.2	1.08	.15	.22	.25	2.22
	3	3.2	1.08	.17	.25	.28	2.46
5	1	1.7	4.34	1.93	--	4.85	4.30
	2	8.6	1.67	.05	--	(.12)	1.70
7		4.2	2.72	.04	--	1.37	3.60

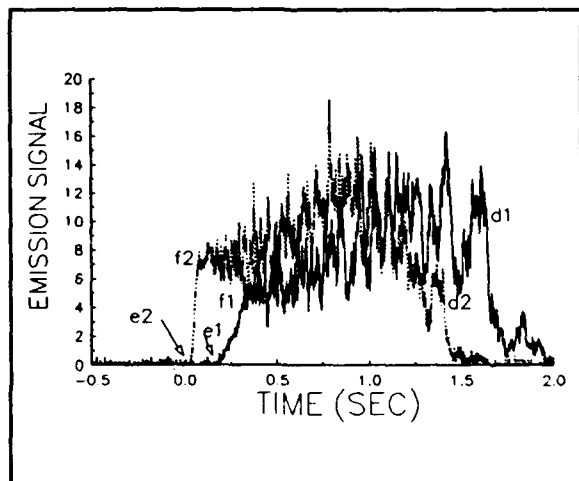


Figure 1. Photomultiplier Record of Emission for M9 in N_2

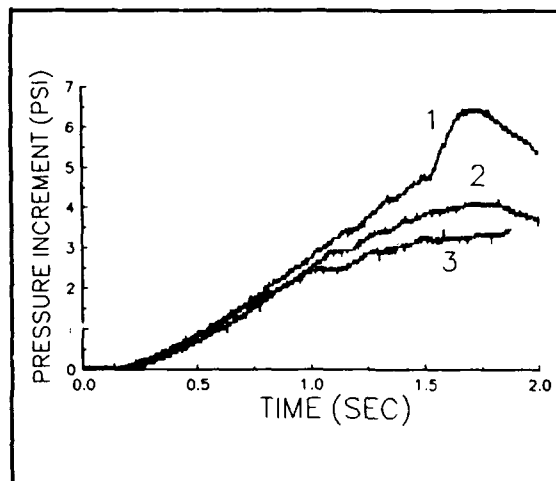


Figure 2. Effect of Oxygen on Pressure for M9.

The effect of flux level on photomultiplier signals during ignition and flamespreading is shown in Figure 1. In this example, the bomb was filled with nitrogen to one atmosphere; the propellant is M9. Both t_e and t_f delays decrease for M9 as power increases, as expected. The M9 emission decay, t_d , is due primarily to consumption of propellant. The effect of air may be important for direct laser ignition of propellant beds since there is always some ullage in gun chambers. To assess its importance requires experiments which mimic the restriction of the air supply by neighboring grains or powder. An attempt was made by surrounding the propellant sample with a cylindrical shroud (plastic, .75 in. ID) with axis coincident with the laser beam, which restricted the path of the incoming air. Figure 2 shows the effect of replacing nitrogen with air and the effect of the shroud on the pressure signals. Signals 1, 2 and 3 correspond to experiments in air, N_2 and air with the shroud in place, respectively. The most obvious effect of the presence of oxygen is the increase in pressure level after deradiation. The temporal behavior is relatively unchanged. The presence of the shroud seems to eliminate the oxygen effect.

The lines in Figure 3 are the calculated ignition delays as a function of flux levels for different heat release rates. "N" refers to a rate determined from conductive (radiant) ignition experiments with the Russian double base propellant N (0.58 NC, 0.28 NG, 0.12 DNT)¹ and "M9" refers to a rate determined from shock tube experiments with M9 (0.58 NC, 0.40 NG) powders in nitrogen ($T = 600-700$ K and pressure = 0.9-3.5 atm).² For N and M9, values of Q_z (cal/g-s) are $2.5E16$ and $2.6E11$, respectively. The corresponding values for E (kcal/mole) are 35 and 22. The rate for N propellant was based on surface thermocouple measurements of delays to

a jump in temperature which presumably involved flamespreading while the rate for M9 was determined from emission (delays) from small ($<10 \mu\text{m}$), uniformly heated particles which react (explode) without flamespreading. Deriving global Arrhenius parameters from these different "ignition" delay data involve different assumptions or models of the heat-up and heat release processes. The rate constants are expected to be valid for thermochemical states obtained under the experimental conditions. The use of M9 and N rates in the calculations are expected to produce ignition delays which correspond to measurements of emission and flamespreading delays, respectively. Although Arrhenius parameter for the two rate constants differ markedly the calculated ignition delays are almost identical for flux values greater than $2 \text{ cal/cm}^2\text{-s}$.

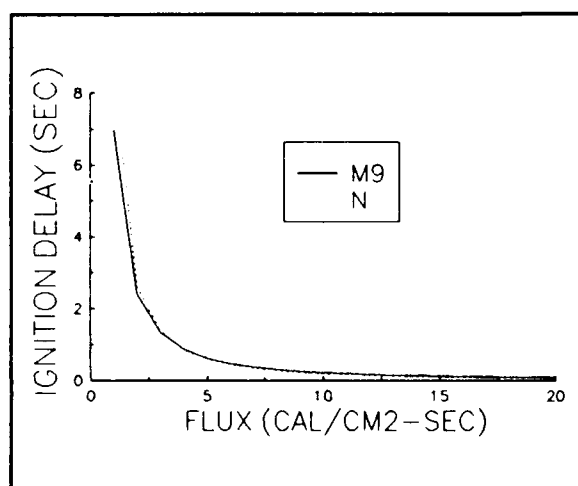


Figure 3. Comparison of Approximate Model Ignition Delays for M9 and N Kinetics.

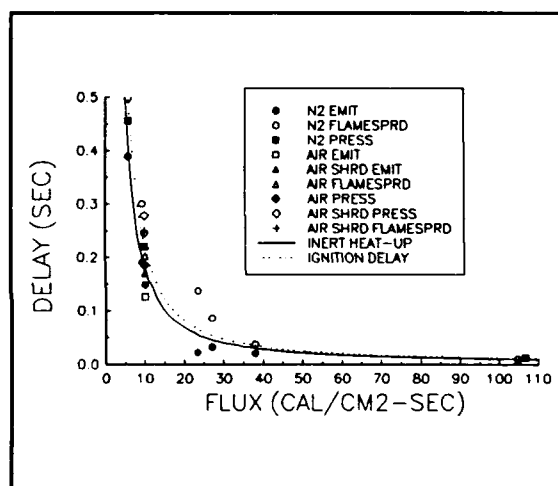


Figure 4. Comparison of Emission and Flamespreading Delays for M9 with Calculations Using Experimental Heat Release Rates.

Figure 4 shows the t_e , t_p and t_f data obtained from experiments with M9 in the bomb under different ambient conditions. The limited amount of data at flux = $10.8 \text{ cal/cm}^2\text{-s}$ (listed in Table I for $p = 3.2 \text{ w}$) suggest that replacement of N_2 by air (with and without a shroud) does not have a great effect on the M9 ignition process. The lines in Figure 4 are the calculated heat-up and ignition delays from the "approximate" model using the M9 kinetics and show that the difference between heat-up and ignition decreases with increasing flux levels. If it is assumed that t_e corresponds to the end of heat-up and that t_p and t_f correspond to ignition, then, considering measurement errors and the uncertainties in thermophysical properties, the predictions of the "approximate" model seem reasonable.

Figure 5 shows the emission signals from experiments with JA2 in argon, initially at ambient conditions, at two flux levels which show the possible effect of deradiation on propellant reaction. The emission (t_e) and flamespreading delays (t_f) decrease with increasing power, similar to the M9 results, but at 8.6 watts the emission decay (t_d) corresponds to laser deradiation (t_l) and propellant consumption is not completed. Emission is due to local combustion in the irradiated region resulting in formation a crater in the unconsumed propellant. At lower power and longer pulse lengths reaction decelerates at t_l (corresponding to d_l in the figure) but recovers. Self acceleration leads to the subsequent sharp increase in emission at (f) which is assumed to be the start of flamespreading. The subsequent decay (not shown) in emission is the result of propellant consumption.

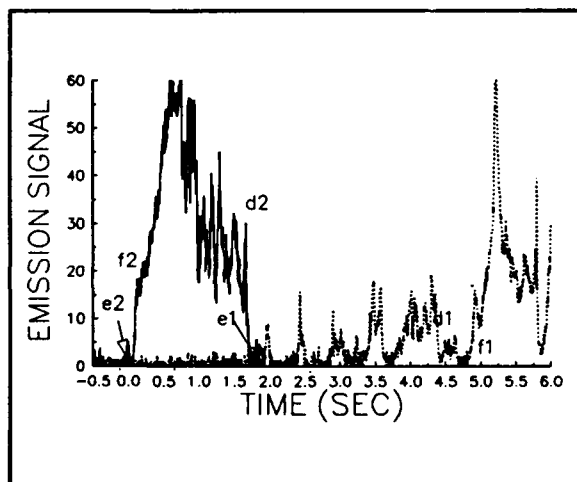


Figure 5. Emission Signals Showing Different Effects of Deradiation on Reaction of JA2 in Ar

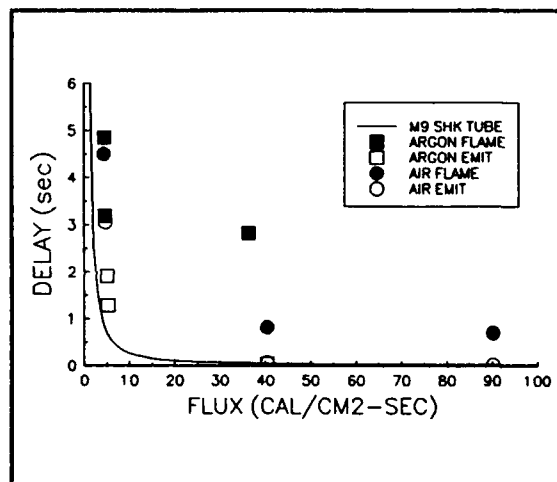


Figure 6. Comparison of Emission and Flamespreading Delays for JA2 in Ar and Air with Calculations

The data in Figure 6 are the emission and flamespreading delays for JA2 in air and Ar initially at ambient conditions. The data at 90 cal/cm²-s are from an experiment in room air. The others were taken in the bomb. The Arrhenius parameters have not yet been determined for JA2 and the line in Figure 6 is the calculated M9 ignition delays shown previously in Figure 4. The data near 40 cal/cm²-s suggest that substitution of air for argon shortens flamespreading delays but does not affect emission delays. The differences between calculated delays and measured emission delays are probably within estimated measurement/thermophysical properties errors; except perhaps, at low flux levels, flamespreading delays are much greater than calculations.

Figure 7 shows the emission signal from ignition of M30 in room air in which a transition from fizz (f) to flame (g) burning occurs. This transition has only been observed during irradiation and differs from the transition to flamespreading shown in Figure 5 (f1) which can occur after deradiation. The data in Figure 8 are the emission and flamespreading delays for M30 in room air and in the bomb with N₂ at ambient conditions. The line is the calculated M9 ignition delay. The results suggest that emission delays are not greatly different in air or N₂ and except at low flux values are relatively constant, which is in reasonable agreement with calculations. The flamespreading delays are less in air than in nitrogen (near 35 cal/cm²-s) and are much greater than calculated values using rates for M9 (and N) propellant.

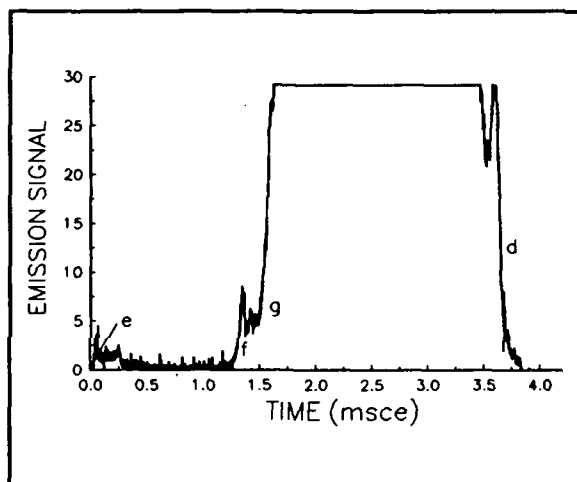


Figure 7. Emission Signal Showing Transition from Fizz to Flame Burning for M30 in Air

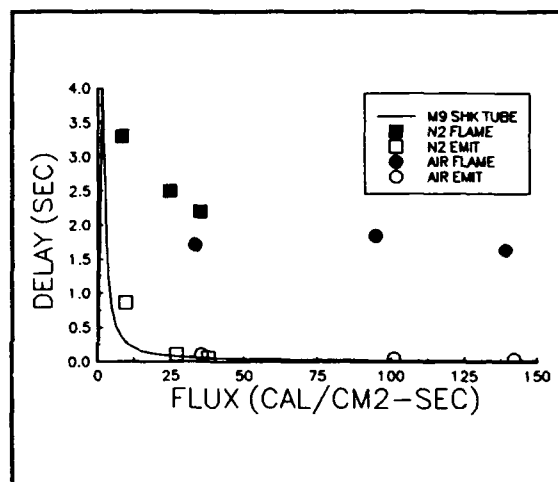


Figure 8. Comparison of Emission and Flamespreading Delays for M30 in N₂ and Air with Calculations.

Comments and Tentative Conclusions

1. Due to space limitations, the results for the nitramine propellants M39 and HMX2⁴ have been omitted. For M39, using rate constants derived from DSC experiments, predictions of the "approximate" model were in agreement with t_g but not t_f data. For HMX2, using rate constants derived from either DSC or shock tube experiments, the model was unable to predict either t_g or t_f data.

2. Observations of deradiation and ambient gas effects similar to present observations but at higher pressures (5-20 atm) were reported in an early investigation

BEYER AND COHEN

of radiative ignition of double-base propellants.³ The present "approximate" model can not be used to predict such effects.

3. The substitution of air for inert gases leaves t_e values unchanged but shortens (except for M9) t_i values. This suggests that reaction of gases (O_2) affect flamespreading but not initiation.

4. The good agreement between emission delays (t_e) for radiatively heated M9 disks and calculated values using heat release rates determined from emission delays for shock heated M9 powder strongly suggests that the inverse process may be possible, i.e., using the "approximate" model to obtain "useful" rate information for propellants with ignition delays not greatly affected by ambient gases.

5. The decrease in the flux dependence of the calculated ignition delay (for all input rates) as flux levels increase suggests that precise knowledge of initiation rates are not critical inputs into models to predict flamespreading at high flux levels.

6. The laser assisted transition from fizz to flame burning with M30 in air (Figure 7) suggests the possibility that laser absorption by reaction products (CO_2) aids in the transition.

7. Data at higher power are necessary to determine usefulness of ignition models for the LIGHT Program.

These comments emphasize the short-comings of the "approximate" model. The inability of this model to predict behavior of many solid propellants which are of interest to the LIGHT program (M30, JA2 and LOVA type) implies that the next level laser ignition models (for many propellants) need to include explicitly the effect of gas phase reactions on the physics of flamespreading.

II. SMALL CHAMBER DIRECT IGNITION

In addition to the studies described above to characterize the laser-propellant interaction, an effort has been made to study the direct laser ignition of propellants. The ready ignition of several propellants in the open air has been earlier demonstrated in our lab and elsewhere. In the present work, experiments have been started to study ignition in confined vessels, which are a step closer to ignition in a gun chamber. This section of the paper discusses primarily the ignition of small samples of solid and liquid propellant that are confined by a chamber with limited ullage. The ignition characteristics are observed and guidelines are set for future work.

BEYER AND COHEN

Experimental

The chamber used in most of these studies is shown in cross section in Figure 9. It was machined out of 316 stainless steel with a chamber capacity of 0.1 cm^3 . The use of the mylar as a combination window and blowout disk has proven to be advantageous for bench top survey studies. The back wall is formed by the pressure gauge, as shown. No protective grease was used on the gauge because of the very limited volume and concerns that the grease would introduce significant non-reproducible effects on the volume and heat transfer to the walls. Thus, some pressure traces have the characteristic negative shift due to heating of the pressure transducer; this shift is only cosmetic for these studies.

The laser used in these studies was a neodymium glass laser operating at $1.054 \text{ }\mu\text{m}$ wavelength with pulse length variable from $100 \text{ }\mu\text{s}$ to 10 ms and variable energy. For these studies the light was

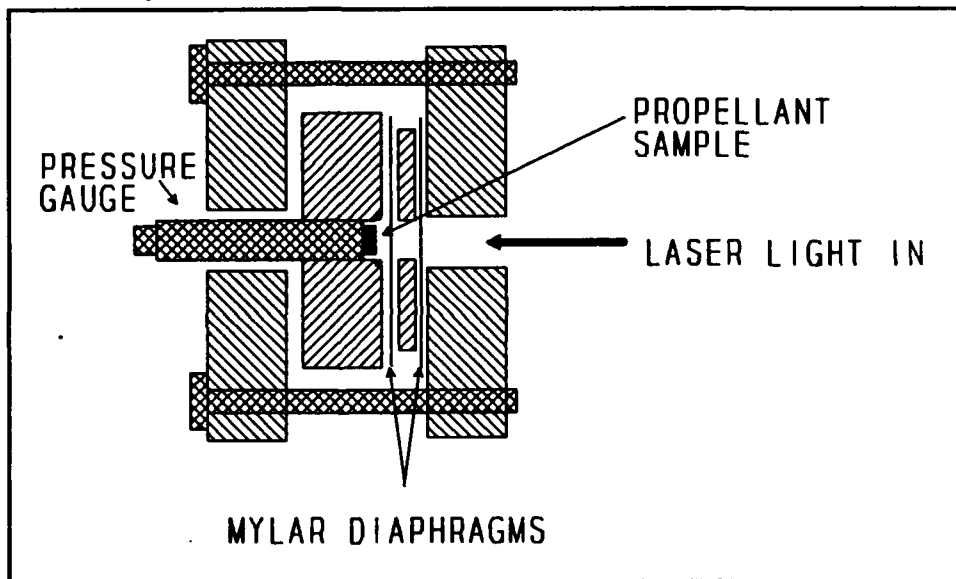


Figure 9. Schematic Diagram of the 0.1 cm^3 Test Chamber.

focused to a 0.125 inch diameter spot with a 2 -inch focal length lens. Except as noted, the mylar was 0.020 inch thick. For some of the studies a second expansion chamber was created by adding a spacer after the diaphragm with a second diaphragm. The volume of this second chamber was about 0.05 cm^3 or five times the first chamber volume. The point of rupture of the first layer diaphragm is obvious from the sharp pressure drop in each figure discussed below. This dual-chamber arrangement allowed the study of how completely the propellant was ignited and whether sudden, but incomplete, depressurization would extinguish it. The results obtained from this configuration will also be important in the study of laser ignition where the irradiated material could be placed in a confined volume to enhance

BEYER AND COHEN

ignition; after ignition the confinement would rupture and the following processes take place in a depressurized environment.

In contrast to the $10.6\mu\text{m}$ wavelength CO_2 laser, most of the propellants are quite transparent near $1\mu\text{m}$. Thus, in most of the observations described here, a graphite coating was required on the propellant surface in order to have sufficient energy absorbed to ignite the sample.

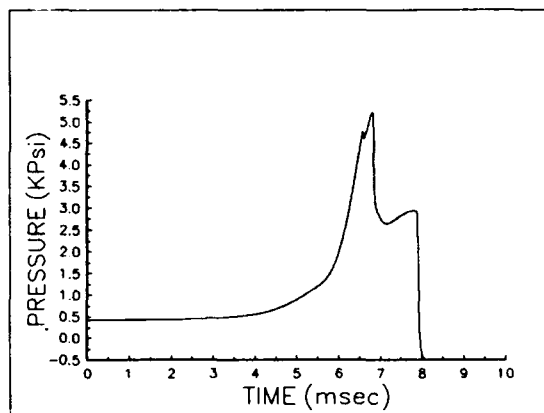


Figure 10. M30 Laser Ignition in Dual Chamber Fixture.

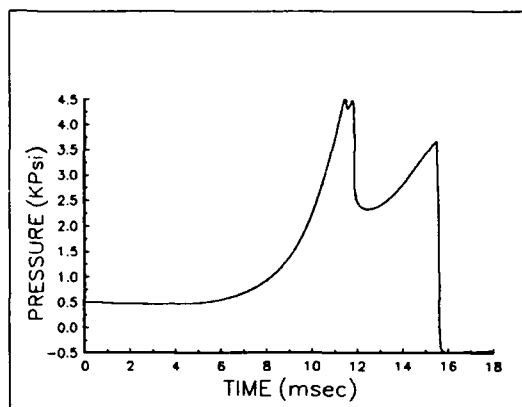


Figure 11. Ignition of M43 Propellant in Dual Chamber Device.

OBSERVATIONS

A. M30 Propellant

The energy to the surface for the M30 propellant ignition event shown in Figure 10 was less than 3J before the apparent ignition time. As can be seen, the second diaphragm broke at lower pressure than the first. The pressure behavior before the burst of the first diaphragm is quite similar to that recorded in the single chamber ignition. In the case of Figure 10, the sample lost 12.1 mg or 9.5 percent of the total mass. The pressure record again shows good sustained ignition. The residual propellant sample from this and similar tests show the smooth glassy surface and enlarged perforations that are indicative of flamespread over the entire surface.

BEYER AND COHEN

B. M43 Propellant

Several LOVA propellants were studied briefly. The response of one of these, M43, is shown in Figure 11. Approximately 8.5 J of laser energy were incident on the solid surface during 10 ms. As can be seen in the plot, the laser raised the pressure sufficiently that full ignition of the propellant took place shortly thereafter. The pressure rise was much slower in comparison with the M30 propellant shown in the previous figure. This result was expected because of the difficulty of igniting the LOVA propellants.

In this case, the original sample was 90.0 mg before irradiation and 86.6 mg after the event shown. The mass lost by a similar piece of material under the same laser conditions but without confinement showed a mass loss of approximately 0.1 mg, which is at our level of measurement uncertainty. The amount of mass lost when an M43 propellant sample was ignited for these brief times varied from 2 to 9 mg and appeared to be independent of the original mass. These values of mass loss, as well as the appearance of the samples after the tests, show that vigorous, complete ignition of the samples has taken place. The variability in sample consumption may be related to either relative flamespread rates or differences in diaphragm rupture behavior.

C. Liquid Propellant 1846

Although the liquid propellants are also quite transparent at $1.054\text{ }\mu\text{m}$, most experiments were done without adding graphite. The mechanism of the energy coupling into the liquid is the subject of continuing investigation. This propellant also ignited vigorously and promptly with the laser source. As can be seen in Figure 12, the ignition delay was about 5 ms for this typical example. The liquid propellant is most obviously different from the solids in that it blew itself out of the chamber almost completely. In the single diaphragm experiments, it behaved much like the M43 in the sense of needing good confinement and pressurization during the laser pulse in order to achieve ignition.

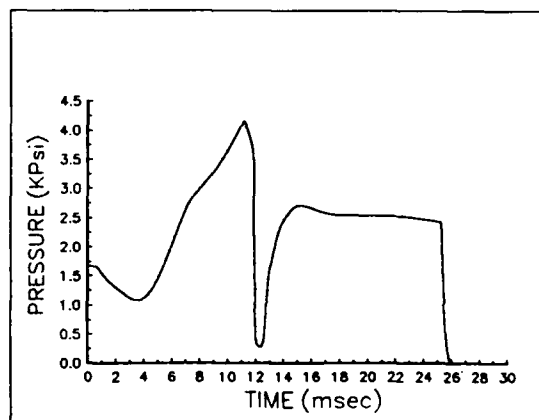


Figure 12. Ignition of LP1846 with Vented Second Chamber.

BEYER AND COHEN

The behavior of LP1846 in the double chamber studies was significantly different from the solid propellants. The main difference is possibly related to the near extinguishment of the LP1846 followed by rapid re-ignition. As can be seen in Figure 12, the second peak is very sharply rising after almost complete depressurization. The experiment shown in Figure 12 incorporated a small (ca. 100 μm diameter) vent hole which allowed sustained combustion without exceeding the pressure capacity of the test chamber.

SUMMARY

In the first part of this study, a comparison of experimental and calculated values for ignition parameters has been presented which shows limited agreement and points the direction for future improvements in this research.

In the related study, a variety of propellants have been shown to be readily ignitable by laser light at a wavelength that is producible in compact, efficient package size. Ignition times vary substantially around the typical values shown here; the major uncontrolled variable is presently thought to be the coupling of the energy into the propellant.

ACKNOWLEDGEMENT

The authors would like to thank J. Deas, C.S. Miser, K.L. McNesby and R.A. Pesce-Rodriguez of the BRL, for help in acquiring experimental data and A.J. Kotlar, also of the BRL, for help with the calculations.

REFERENCES

1. V.H. Vilyunov and V.E. Zarko, *Ignition of Solids* (Elsevier, 1989).
2. A. Cohen and H.E. Holmes, "Convective Ignition of Double-Base Propellants", 19th Symposium (International) on Combustion, p. 691, The Combustion Institute, 1982.
3. (a) L. DeLuca, T.J. Ohlemiller, L.H. Caveny and M. Summerfield, *AIAA Journal* 14, 7, 940-946, (July 1976). (b) L. DeLuca, T.J. Ohlemiller, L.H. Caveny and L. Summerfield, *AIAA Journal* 14, 8, 1111-1117, (August 1976).
4. A. Cohen and R.A. Beyer, CPIA Pub. 573, Vol. I, pp. 435-440, 1991.

BINDER, MAMMONE, WADE, RONDEAU & LAVENE

Method for Increasing Breakdown Voltages
in Polymer Film Capacitors

Michael Binder, Dr., Robert J. Mammone, Dr.,
and William L. Wade Jr.
U.S. Army Electronics Technology and Devices Lab
Fort Monmouth, NJ 07703-5000
and
Ernest Rondeau and Bernard Lavene
Electronics Concepts Inc.
5260 Industrial Way West
Eatontown, NJ 07724

INTRODUCTION

Future industrial and fielded energy storage applications will require high quality, high reliability capacitors. The maximum electrostatic energy density that can be stored in spirally wound capacitors depends on the product of total capacitance of the capacitor and the **square** of the maximum voltage that can be applied across the capacitor. If overall breakdown strengths, V_b , expressed in volts/micron of thin (8-12 micron) dielectric films can be increased, capacitors constructed from these films can operate at higher voltages which ultimately translates into higher electrostatic energy densities. Organic polymers with high resistivity, low dissipation factors and high electric field breakdown strengths have important applications as dielectrics in wound film capacitors. Surface characteristics of polymer films are important in determining interfacial properties and consequently, the technological performance of these polymers. Since it is difficult to find polymers having both appropriate bulk and surface properties, it is frequently necessary to create newly functionalized surfaces with specific characteristics while retaining desirable bulk properties.

Among various surface modification methods, low temperature gas plasma techniques are inexpensive and relatively easy to use without producing severe thermal damage to the bulk. These plasmas arise when low pressure gases at room temperature are subjected to electrical discharge, e.g., a radio frequency or microwave field. A fraction of the gas is excited by the discharge and form free electrons, ions, radicals, excited molecules, and photons. These activated species can interact with exposed surfaces of materials in the plasma in different ways.

Resulting surface properties of polymers following exposure to gas plasmas depend on the chemical structure of the polymer, type of gas plasma used, treatment time, and parameters associated with the equipment such as gas flow, energy input and geometry. The purpose of the present study was to see whether performance improvements could be observed after exposing films and/or even fully constructed, spirally wound, film capacitors to gas plasma.

EXPERIMENTAL

Six thermoplastic polymer films were studied: polyethylene, polypropylene, polyester, polysulfone, polycarbonate and polyvinylidene fluoride (PVDF). Coupons (5 cm X 17 cm) cut from a roll were exposed for four minutes to either oxygen, helium or $96\%CF_4/4\%O_2$ gas plasmas at power densities of 0.002 watts/cm^2 . A Rame-Hart Telescopic Goniometer was used to measure static contact angles made by water drops on unmetallized films. Aluminum adhesion was measured on films that had been metallized by vapor deposition. Aluminum deposition on our film samples were done by taping polymer coupons onto a polyethylene carrier roll and sending the whole carrier roll (with the taped on samples) through a commercial metallizer where approximately 100-150 angstroms of aluminum was deposited. Elapsed time between exposure to plasma and metallization was approximately one month. Adhesion of aluminum to these polymers was measured by placing a piece of 810 Scotch tape over metallized polymer, removing the tape, and qualitatively determining how much aluminum was removed. Breakdown voltages were measured by holding unmetallized films between 0.635 cm. diameter brass electrodes in silicon transformer oil at room temperature and ramping applied voltage across these electrodes at 500 volts/sec. from zero volts until breakdown occurred and the sample could not hold additional voltage. Spirally wound capacitors were constructed by sandwiching films of polycarbonate (5 microns thick), polypropylene (4 microns thick), or polyester (3.5 microns thick) between 5 micron thick tin/lead foil and winding this around a thin mandrel. Since the width of metal foil was .635 cm. less than that of the dielectric polymer strip, the .031 cm. margin on each side acted as an apron to prevent flashovers.

A total of 36 identical capacitors of each type was constructed and divided into three sets. One set of 12 capacitors was used as the control group, one set was exposed to O_2 , and one set was exposed to 96% CF_4 /4% O_2 plasma. All of these capacitors were then subjected to a 16 hour bakeout at 54°C. Breakdown voltages were measured in air at room temperature.

TABLE 1. Contact angle of water (in degrees), relative adhesion of vapor deposited aluminum (described as either A: excellent, B: good, or C: poor), and breakdown strengths in volts/micron of various polymers that had been exposed to either oxygen, CF_4/O_2 or helium gas plasma. For each polymer and gas plasma studied, contact angles are listed on the upper left, relative adhesion of aluminum is listed on the upper right and breakdown strength, V_b , is listed in parentheses below the line.

	Unexposed	Exposed to oxygen plasma	Exposed to CF_4/O_2 plasma	Exposed to helium plasma

Polycarbonate	<u>72/B</u> (724)	<u>39/A</u> (752)	<u><15/A</u> (764)	<u>37/A</u> (760)
Polysulfone	<u>70/A</u> (445)	<u>25/A</u> (472)	<u><15/A</u> (484)	<u>26/A</u> (484)
Polyester	<u>66/C</u> (681)	<u>29/A</u> (669)	<u>30/A</u> (673)	<u>29/A</u> (677)
PVDF	<u>71/C</u> (598)	<u>40/A</u> (630)	<u>70/A</u> (661)	<u>57/A</u> (641)
Polypropylene	<u>98/C</u> (709)	<u>40/A</u> (701)	<u>72/A</u> (744)	<u>53/A</u> (732)
Polyethylene	<u>90/C</u>	----	<u>20/A</u>	<u>50/A</u>
Polystyrene	<u>83/B</u> (480)	<u>15/A</u> (520)	<u><15/A</u> (547)	<u>26/A</u> (512)

EXPERIMENTAL RESULTS

WATER WETTABILITY OF POLYMER FILMS

Table 1. summarizes our experimental results on water wettability of these polymers after they had been exposed to each of the three gas plasmas. For each polymer and gas plasma, water contact angle in degrees is listed on the upper left side of the line. For unexposed polymers, contact angles range from 65-95 degrees which is typical for hydrophobic surfaces. Contact angles in almost all cases, decreased dramatically after these polymers had been exposed to any of the gas plasmas. Decreased contact angles indicated that polymer surfaces had become more receptive to water (more hydrophilic).

Exposure to helium plasma had the least effect on contact angle for all of the polymers tested (with the exception of polyester). Exposure to oxygen plasma had the greatest effect in reducing contact angle for PVDF and polypropylene, while exposure to CF_4/O_2 plasma had the greatest effect in reducing contact angles for polycarbonate, polystyrene and polysulfone.

ADHESION OF ALUMINUM TO POLYMER FILMS

Table 1. also lists (on the upper right side of the line) the relative adhesion of vapor deposited aluminum to each polymer. Here, designation A indicates excellent adhesion of aluminum, designation B indicates good adhesion of aluminum, while designation C, indicates poor adhesion of aluminum. Clearly, aluminum adhesion to polycarbonate, polyester, PVDF, polyethylene, and polypropylene (for which adhesion to untreated samples is poor) improved substantially after exposure to gas plasma. In fact, all polymer samples tested had excellent adhesion of aluminum after being exposed to any of the three gas plasmas. Since one month had elapsed between the time that these polymers had been exposed to gas plasma and the time that they were metallized, effects of exposure to gas plasma on adhesion of vapor deposited aluminum is apparently retained even after these polymers were exposed to ambient atmosphere for one month.

BREAKDOWN VOLTAGE OF POLYMER FILMS

Table 1. also lists DC V_b for these polymers before and after exposure to various gas plasmas. For polysulfone, polystyrene, and PVDF films, V_b increased following exposure to selected gas plasmas. Measured V_b of PVDF films for example, increased by about 11% following exposure to either CF_4/O_2 or helium plasmas while V_b for polystyrene increased by about 14% following exposure to CF_4/O_2 plasma.

TABLE 2. Overall breakdown voltages, in volts, and total capacitance measured at 1 kHz, of fully constructed, loosely wound polycarbonate, polyester and polypropylene based film capacitors before and after they had been briefly exposed to either oxygen or CF_4/O_2 gas plasma. "Low" and "high" refer to the lowest and highest breakdown voltages measured within a given set.

CAPACITOR TYPE	<u>BREAKDOWN VOLTAGE,</u>			<u>CAPACITANCE</u>
	<u>VOLTS</u>			@ 1kHz.,
	<u>LOW</u>	<u>HIGH</u>	<u>AVERAGE</u>	mF
POLYCARBONATE				
unexposed	96	378	276	0.179
exposed to O ₂ plasma	211	406	310	0.183
exposed to CF ₄ /O ₂ plasma	343	869	555	0.182
POLYESTER				
unexposed	380	1280	879	0.341
exposed to O ₂ plasma	310	1280	880	0.344
exposed to CF ₄ /O ₂ plasma	770	1290	1083	0.342
POLYPROPYLENE				
unexposed	590	1060	809	0.205
exposed to O ₂ plasma	640	860	741	0.182
exposed to CF ₄ /O ₂ plasma	410	1110	841	0.185

BREAKDOWN VOLTAGES OF CAPACITORS

Overall breakdown voltages of loosely wound, fully constructed polycarbonate, polyester and polypropylene based capacitors are listed in Table 2. Exposure of these capacitors to 96%CF₄/4%O₂ plasma produced more dramatic increases in breakdown voltages than did exposure to O₂ plasma. For polycarbonate based capacitors, brief exposure to 96%CF₄/4%O₂ plasma **doubled the breakdown voltages** as compared with baseline capacitors! Polyester based capacitors showed a more modest 23% increase in breakdown voltages after exposure to 96%CF₄/4%O₂ plasma while polypropylene based capacitors showed only 4% improvement in breakdown voltage following exposure to CF₄/O₂ plasma.

DISCUSSIONS

Observed changes in adhesion of aluminum, contact angle of water and breakdown strengths of thin polymer films following exposure to gas plasma may be simply attributed to removal of normally adsorbed impurities from polymer film surfaces. However, a more plausible explanation is that exposure of polymer surfaces to gas plasma forms reactive chemical groups, which, upon subsequent exposure to the atmospheric oxygen, may allow covalent oxygen bonds to form. Vapor deposited aluminum can react with these hydrophilic oxygen groups to form strong chemical bonds. This would account for improved aluminum adhesion in polymers that had been exposed to gas plasmas. Chemically modified surface regions could also trap or impede injected charge from entering the bulk thereby decreasing leakage currents and increasing V_b.

Exposing PVDF to CF₄/O₂ plasma, for example, may be replacing hydrogen atoms with fluorine to form a modified, highly resistive, fluorinated surface region. This layer may inhibit charge injection and thereby increase V_b. This explanation is somewhat supported by the fact that water contact angle decreased the least for PVDF following exposure to CF₄/O₂. The relatively large contact angle measured on PVDF even after exposure to CF₄/O₂ plasma suggests that formed surfaces are still hydrophobic (which would be true for fluorinated layers).

The large increase in overall breakdown voltage observed in polycarbonate based capacitors after exposure to CF₄/O₂ plasma is surprising and unexpected. It can not simply be due to moisture removal because all of these capacitors (including untreated samples) were subjected to similar extensive bakeout procedures. Likewise, we cannot

assume that the work function of tin/lead foil is increased after exposure to CF_4/O_2 plasma, since, if this were the case, one would also expect substantially increased overall breakdown voltage for spirally wound polypropylene or polyester capacitors which were also constructed with the same tin/lead foil.

Post mortem analysis on these capacitors after failure (due to breakdown) showed that most pinholes where breakdown had taken place were located in the region where polymer film was pressed against the edge of the metal foil. Here, electric field stresses combined with mechanical deformation of the polymer film at metal foil edges, weaken the film and make it easier for breakdown to occur. Perhaps gas plasma permeates between the film foil windings and modifies film edges by providing a highly insulating coating which prevents arcing or polymer breakdown. This hypothesis is somewhat supported by observation of water wettability. The water contact angle in the center portion of the polycarbonate film (that had been covered by the tin foil) was approximately the same as that for unexposed polycarbonate film. However, near film edges, where metal foil had not covered the polycarbonate film, the contact angle appeared to be considerably lower. This observation suggested that CF_4/O_2 plasma had indeed permeated the core windings (at least up to the embedded metal foil). However, we still cannot explain why polycarbonate-based capacitors show such huge improvements in breakdown voltages over the other two capacitor types. Perhaps chemical impurities or structural imperfections commonly present near surfaces of solution cast polycarbonate film are modified and/or removed by reaction with CF_4/O_2 plasma. Once these low breakdown species are eliminated, overall capacitor breakdown voltages are enhanced. Capacitors constructed from polypropylene or polyester films (which are not formed by solution casting) do not exhibit such dramatic enhancement after exposure to plasma.

Our research efforts are still at a formative stage and no attempt has been made here to optimize plasma treatment conditions. Consequently, perhaps even greater improvements than those illustrated here may be achievable by "optimized" plasma treatments through variation in the plasma gas used, applied power densities, exposure times or temperatures. Such studies are underway.

BINDER, MAMMONE, WADE, RONDEAU & LAVENE

SUMMARY

Experimental results demonstrate a distinct advantage of exposing polymer films and even fully constructed spirally wound film/foil capacitors to low temperature gas plasmas. Improvements in adhesion of vapor deposited aluminum, increased water wettability, and in some cases, improvements in dielectric breakdown strength or, in fully constructed capacitors, increased overall breakdown voltage were observed after these films or capacitors were briefly exposed to gas plasmas.

ACKNOWLEDGEMENTS

This work was performed under a Cooperative R&D Agreement between Electronics Concepts Inc. and ET&DL. We thank Bernie Rapp from Gasonics Inc. for his invaluable help in treating our samples with gas plasmas and for many useful technical discussions. We thank Ron Siegler of Electronics Concepts Inc. for expertly winding these test capacitors.

BISHOP, SERGI, EIDELMAN, LOTTATI

Development of the Mine Resistant Vehicle Using
Advanced Computational Methods

Mr. Steven S. Bishop* and Mr. Sergio A. Sergi
U.S. Army Belvoir, Research, Development and Engineering
Center (BRDEC), Fort Belvoir, VA 22060-5606
Dr. Shmuel Eidelman and Dr. Isaac Lottati
Science Applications International Corporation (SAIC),
McLean, VA 22102

INTRODUCTION

Today, logistic and combat service support vehicles must operate in a variety of environments. Low intensity conflicts are a continuing threat in many regions around the world, and protecting military vehicles from the effects of land mine detonation is of considerable interest. Our study uses advanced computational methods to study the interaction of the 5-ton M925A1 cargo truck with a land mine. The truck's flat cargo bed measures 4.2 meters long, and 2.4 meters wide, and rests on a chassis positioned about 1.4 meters above the road. The forward cab rests 1.4 meters above the road. The considerable standoff distance of the chassis and the cab offers a natural protection from the mine blast effects. In this paper, we will demonstrate the use of Computational Fluid Dynamics (CFD) methods for a detailed study of a land mine blast and M925A1 truck cross section interaction.

BATTLE DAMAGE DATA OF LAND MINE INCIDENTS

Data for land mine incidents are limited since recording mine encounters is not a high priority during wartime. The primary sources of information to date have been the Vietnam conflict and the Rhodesian bush war. A second limiting factor in data acquisition is that a battle damage assessment reporting team usually is not operational for an entire conflict or war. The Survivability & Vulnerability Information and Analysis Center (SURVIAC), Wright-Patterson Air Force Base, OH, has compiled a database based on Battle Damage Assessment Reporting Team findings. The available information from Southeast Asia dating from April 1969 to June 1970 is summarized for mine incidents. In this database, there are 547 incidents recorded with various vehicle types, including 40 incidents of wheeled vehicle mine

BISHOP, SERGI, EIDELMAN, LOTTATI

encounters. The vehicle types range from a .25 to a 5 ton capacity wrecker. Fifteen of the recorded incidents involve a 5 ton truck. The data also report that 39 incidents, an overwhelming majority, were from a standard or conventional land mine. One incident resulted from an improvised mine. Of the 39 vehicles, 34 were rendered inoperable; 5 vehicles were scrapped.

From a structural standpoint, the location of "hit" or the position of the mine's detonation as it strikes the vehicle is significant. The following breakdown shows that the most common location of the land mine incidents for a wheeled vehicle is with a tire.

<u>Number of Incidents:</u>	<u>Hit Location:</u>
4	Body
1	Body/Wheel
2	Track
32	Wheel
1	Wheel/Fuel Tank

APPROACH AND METHODOLOGY

In order to improve the land mine resistance of the M925A1 truck, first we need to understand the details of the blast wave structure interaction. By doing so, we can estimate the extent of required changes to the truck configuration for various levels of protection. The primary goal is to assure sufficient protection for the driver and passengers to survive a single incident of mine detonation, with minimal modification to the truck geometry.

To study the details of mine blast effect on the truck, we use advanced methods of CFD represented in the Adaptive Unstructured Grid Upwind Second order on Triangles (AUGUST) code developed by SAIC. This code allows detailed description of the truck geometry and high resolution simulation of shock wave structure interaction. Using this code, we can accurately predict the pressure and impulse that will affect various areas of the cargo bed or cabin as a result of mine detonation and assess survivability for the personnel. We also can use the pressure/history data from the blast/structure interaction as an input to the available structural analysis program (ANSYS or DYNA) and assess the damage to the trucks inflicted by the mine detonation.

The accuracy of the data provided by the CFD simulation is very important to calculations of the structural response. However, these two problems can be decoupled in time and the CFD simulation can proceed on the assumption that the structure is rigid. This is possible because the duration of the gas dynamic phase of the blast structure interaction is about 2 milliseconds (msec) and the structural response phase starts 10 to 20 msec after the detonation.

UNSTRUCTURED TRIANGULAR GRID

Practical numerical simulations of fluid dynamic problems call for modeling flows over complicated shapes. In addition, important flow features such as shed vortices, shock waves, slip lines, and boundary layers usually have widely varied lengths and time scale, and need to be resolved. Accurate solution of these problems requires computational grids dynamically adapted to the evolving flow feature, and with full control over solution accuracy in the key regions of the computational domain. It is commonly accepted that only unstructured grids can provide full flexibility in obtaining the local grid resolution sufficient to accurately resolve subscale flow features. The five years since CFD research introduced these grids and methods have produced landmark simulations clearly demonstrating their advantages.^{1, 2, 3}

The Second Order Godunov method was implemented on rectangular grids⁴ and in a few cases on general quadrilateral grids.⁵ This approach has limited application, since structured quadrilateral grids have great difficulty describing a complicated computational domain with multiple bodies of different geometries and scales. Recently, we implemented the Second Order Godunov for unstructured triangular grids.⁶ This enables us to combine the robust and accurate numerical algorithm with a gridding technique, which allows us to describe very complex domains with ease and efficiency. In addition, we developed a novel dynamic grid adaptation methodology that only allocates a dense computational grid to regions where enhanced resolution is needed to resolve strong gradients in flow parameters. As demonstrated herein, this enables an extremely economical allocation of computational resources, and accurate simulation of a complicated phenomenon that includes multiple reflection and interaction of the shock waves.

MATHEMATICAL MODEL AND INTEGRATION ALGORITHM

We consider a system of two-dimensional Euler equations written in conservation law form as:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \quad (1)$$

where

$$U = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ e \end{bmatrix}, \quad F = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(e + p) \end{bmatrix}, \quad G = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(e + p) \end{bmatrix}$$

Here u, v are the x, y velocity vector components, p is the pressure, ρ is the density, and e is total energy of the fluid. We assume that the fluid is an ideal gas and the pressure is given by the equation-of-state

$$p = (\gamma - 1) \left(e - \frac{\rho}{2} (u^2 + v^2) \right) \quad (2)$$

where γ is the ratio of specific heats. It is assumed that an initial distribution of the fluid parameters is given at $t = 0$, and the boundary conditions defining a unique solution are specified for the computational domain.

The system of governing equations in Eq. (1) can be written as

$$\frac{\partial U}{\partial t} + \nabla \cdot Q = 0 \quad (3)$$

where Q represents the convective flux vector. By integrating Eq. (3) over space and using Gauss' theorem, the following expression is obtained

$$\frac{\partial}{\partial t} \int_{\Omega} U da + \oint_{\partial\Omega} Q \cdot n dL = 0 \quad (4)$$

where $dL = n dL$, n is the unit normal vector in the outward direction, and dL is a unit length on the boundary of the domain. The variable Ω is the domain of computation and $\partial\Omega$ is the circumference boundary of this domain.

Eq. (4) can be discretized for each element (cell) in the domain

$$\frac{(U_i^{n+1} - U_i^n)}{\Delta t} A_i = \sum_{j=1}^M Q_j^n n_j \Delta L_j \quad (5)$$

where A_i is the area of the cell; Δt is the marching time step; U_i^{n+1} and U_i^n are the primitive variables at the center of the cell at time n and at the update $n+1$ time step; Q_j is the value of the fluxes across the M boundaries on the circumference of the cell where n_j is the unit normal vector to the boundary edge j , and ΔL_j is the length of the boundary edge j . The fluxes Q_j^n are computed applying the Second Order Godunov algorithm, and Eq. (5) is used to update the physical primitive variables U_i according to computed fluxes for each marching time step Δt . The marching time step is subjected to the CFL (Courant-Fredrichs-Lewy) constraint.

We seek a solution to the system of Eq. (1) in the computational domain, which is decomposed into triangles with arbitrary connectivity. An example of the unstructured triangular grid is shown in Figure 1. The grid is used to describe the computational domain around a cross section of an M925 cargo truck and allows detailed description of a complex computational domain.

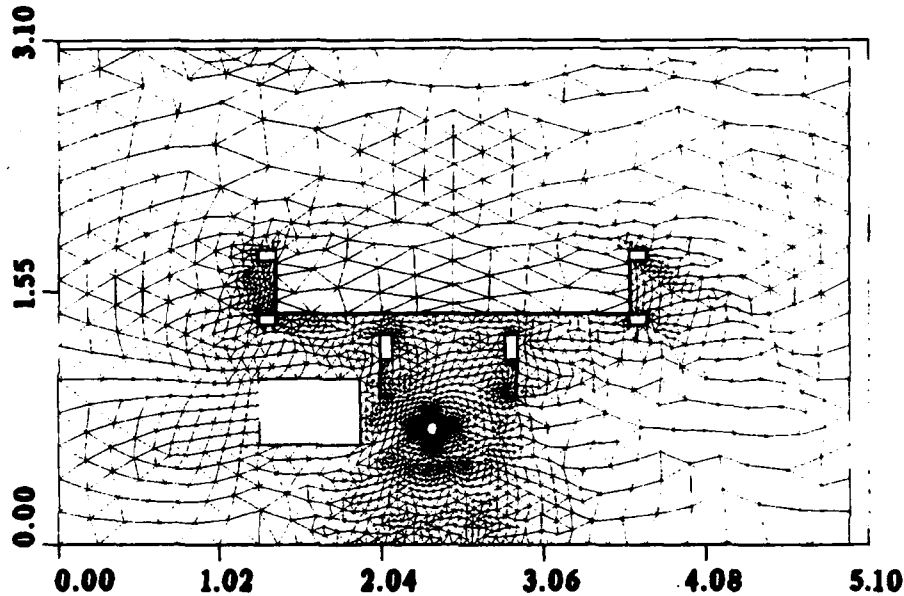


Figure 1. Domain decomposition using unstructured triangular grid.
Grid over a cross section of M925A1 cargo truck.

The numerical technique for solving Euler's equation on an unstructured grid is described in Refs. 6, 7. This numerical technique applies to some of the ideas that were introduced in Refs. 5, 8. The unstructured codes apply the center-based formulation, i.e., the primitive variables are defined in the center of the cell, which makes the cell the integration volume, while the fluxes are computed across the edges of the cell. The basic algorithmic steps of the Second Order Godunov method that is used in our code can be defined as follows:

1. Find the value of the gradient at the baricenter of the cell for each gas dynamic parameter U ;
2. Find the interpolated values of U at the edges of the cell using the gradient values;
3. Limit these interpolated values based on the monotonicity condition;⁷
4. Subject the projected values to the characteristic's constraints;⁹
5. Solve the Riemann problem applying the projected values at the two sides of the edges;
6. Update the gas dynamic parameter U according to the conservation Eq. (1) applying to the fluxes computed and the current time step.

CODE VALIDATION PROBLEM

To validate the AUGUST code, we simulated the problem of interaction of Mach 2.85 planar shock wave, propagating in a channel with a 45° symmetrical double ramp. The results of the simulation will be compared to the experimental and computational results presented in Ref. 9. Figure 2 shows the experimental interferogram of the problem to be simulated. We will present a number of density contour plots showing the flow evolutions in the course of this transient simulation. The ambient pressure and density were set to a unit value and the Mach number is 2.85. The flow in the channel is from left to right.

Figure 2 is the experimental interferogram of the stage of shock/ramp interaction when the shock wave has just passed the ramp. In Figure 2, we observe a complex flow pattern created by the shock ramp interaction. The flow field contains reflected shocks, recompression shocks, four triple points, strong rarefaction fan, slip lines, and vortices. To simulate accurately all the details of the shock/ramp interaction, the computer code should have high resolution for modeling flow regions with sharp gradients.



Figure 2. Interferogram of $M = 2.85$ shock interaction with a 45° double ramp.⁹

Figure 3 shows the computational domain for the shock ramp interaction problem. The computational domain is decomposed by an unstructured triangular grid. The grid is adapted in the area of the shock front, and one can observe a significant increase of the grid density in this area. The area of a triangle in the dense region of the grid is about 100 times smaller than the area of a triangle in the initial grid. In Figure 3, the solution in the form of density contours is shown at the time just prior to the shock impinging on the ramp. Here it is possible to observe that the grid adaptation algorithm adds cells in front of the shock discontinuity and removes them behind the discontinuity; thus, only the area that requires high resolution is covered by the dense grid.

This methodology minimizes the use of computer resources for a required resolution. The adaptive unstructured grid methodology allows high resolution representation of the flow gradients and in Figure 3 we can observe a very sharp planar shock.

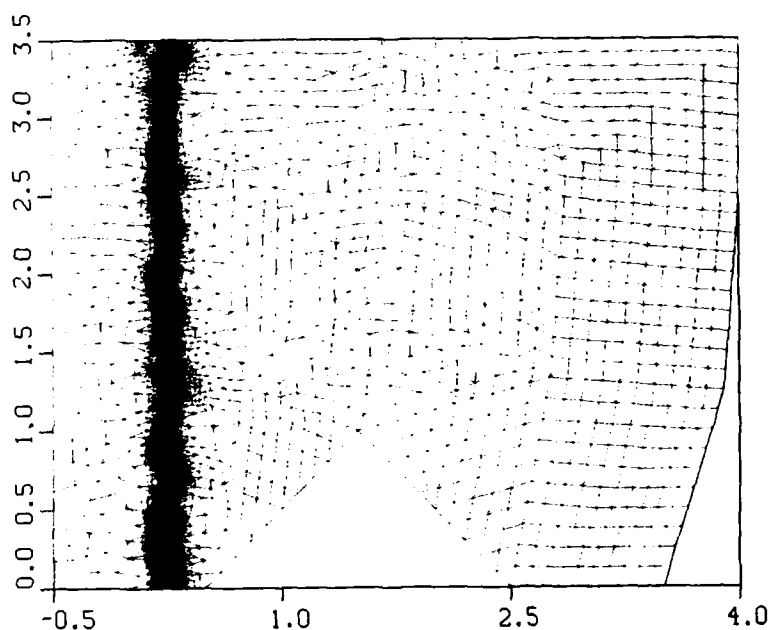


Figure 3. Unstructured grid and shock position before the shock-double ramp interaction. $M = 2.85$. Total number of vertices = 12,599.

In Figures 4 and 5, we show the computational grid and the density contours for the solution of the stage at which the main shock front just passes over the apex of the ramp. The density contours overlay the grid in Figure 4 to illustrate the grid adaptivity. For clarity, contour plots are displayed separately from the grid in Figure 5. Figure 4 shows that the grid adaptivity criteria is successful in capturing the flow gradient. This includes shock waves, contact discontinuities, and slip lines. In Figure 5, we can see very sharp resolution of the complex flow structure. Triple point of the Mach reflection, slip line, and strong vortex forming as a result of strong expansion around the apex are clearly visible.

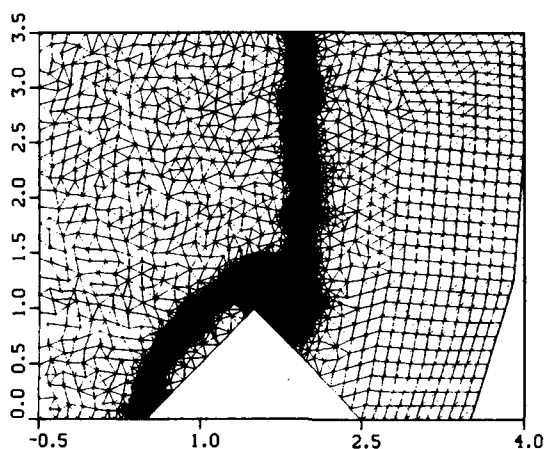


Figure 4. Unstructured grid for the shock-double ramp interaction after the shock passed the apex. $M = 2.85$. Total number of vertices = 22,126.

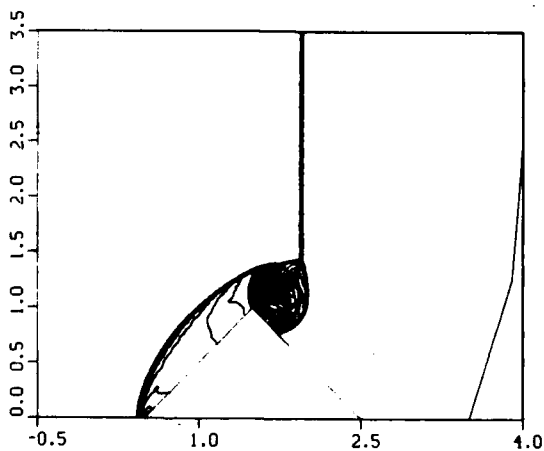


Figure 5. Density contours for the stage shown in Figure 4. $\rho_{\max} = 7.593$; $\rho_{\min} = 0.096$; $\Delta\rho = 0.214$.

In Figures 6 and 7, the simulation results are shown at the stage that corresponds to the time of the experimental results shown in Figure 2. We can observe that the grid is adapting to a changing flow pattern. All the elements of the complex flow are captured by the dynamically adapting grid. The over-expansion around the ramp apex led to formation of the recompression shock, which follows the strong vortex originated at the apex. We also can observe a formation of the triangle of shock waves that originates a vertical slip line that follows the main shock which is typical for this interaction. All the flow features of the experimental interferogram can be identified in our simulation result shown in Figure 7. Further analysis of the values of the contour levels and the density levels shown in the interferogram validate a numerical result accuracy of more than 4%, which is well within the accuracy of the experimental result.

In addition to the results shown in Figures 3 through 7, we have validated the AUGUST code for a range of supersonic, transonic, and subsonic transient and stationary flow problems. The results presented in our previous publications^{5, 6, 7} consistently demonstrate the high resolution and accuracy capabilities of the code, and the ability of the adaptive grid technique to capture the details of complex flow regimes.

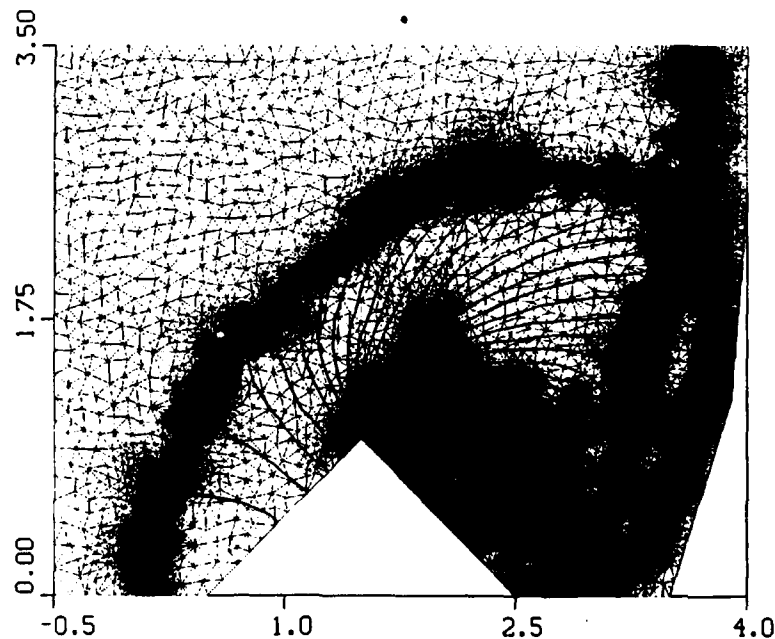


Figure 6. Unstructured grid for the shock position at the stage shown in the experimental interferogram. $M = 2.85$. Total number of vertices = 79,035.

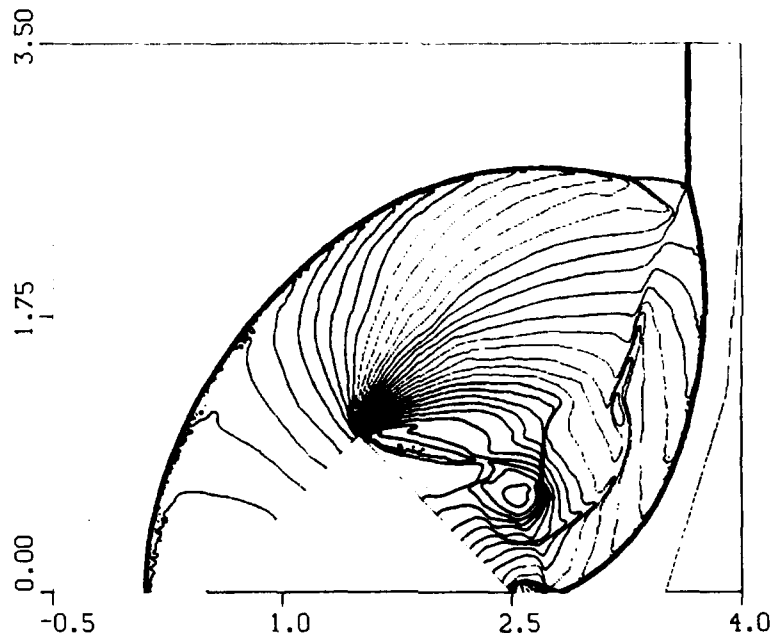


Figure 7. Density contours for the stage shown in Figure 6
 $\rho_{\text{max}} = 7.622$; $\rho_{\text{min}} = 0.082$; $\Delta\rho = 0.215$.

TRUCK-LAND MINE BLAST INTERACTION SIMULATION

We present here the results for the base case of the mine resistant vehicle study: detonation of the mine triggered by the wheels of the unprotected truck. The truck cross section geometry (Figure 1) includes cargo bed, support beams, fuel tank, and the drive shaft. We can observe that the unstructured grid allows accurate description of the complex geometry having elements of different length scale. We have selected this case as a reference case because it will allow us to test our simulation methodology and give data of pressures and impulses that are typical for the unprotected M925A1 truck. We simulate interactions of the cross section shown in Figure 1 with the blast generated by the detonation of a single M-15 AT mine (about 20 pounds of TNT).

The planar two-dimensional presentation of the truck cross section will represent the worst case scenario, because of lack of expansion into the third dimension, and will be typical for the area in the middle of the cargo bed. To take into account three-dimensional expansion of the detonation product from the land mine detonation, we have simulated the initial stages of the blast propagation using two-dimensional axisymmetric formulation. This simulation was done by a structured code that also takes into account the equation-of-state for the detonation products. This code, SOG2DSE (for Second Order Godunov 2 Dimensional for Solid

Explosives), was developed by SAIC and is currently used by BRDEC for predicting the effects of mine detonations. Two-dimensional axisymmetric simulation will give an accurate account for the three-dimensional effects for a symmetric blast propagation. The axisymmetric simulation was carried out up to the point at which the incident shock encounters the elements of the truck geometry. At this point, the result of the axisymmetric simulation on structured grid is mapped up to unstructured triangular grids. Thus, the simulation on unstructured grid starts at the stage shown in Figure 8, where the results are represented by pressure contours in pascals. Here, the maximal pressure is observed in the region of detonation products and is ≈ 160 MPa. To accommodate the high gradients of pressure, the grid density is significantly increased in the area of the detonation products. The initial grid shown in Figure 1 had 2,005 vertices; the grid at the stage shown in Figure 8 has 7,873 vertices. The 5,800 additional vertices were used to cover the region of the propagating blast wave to assure high resolution.

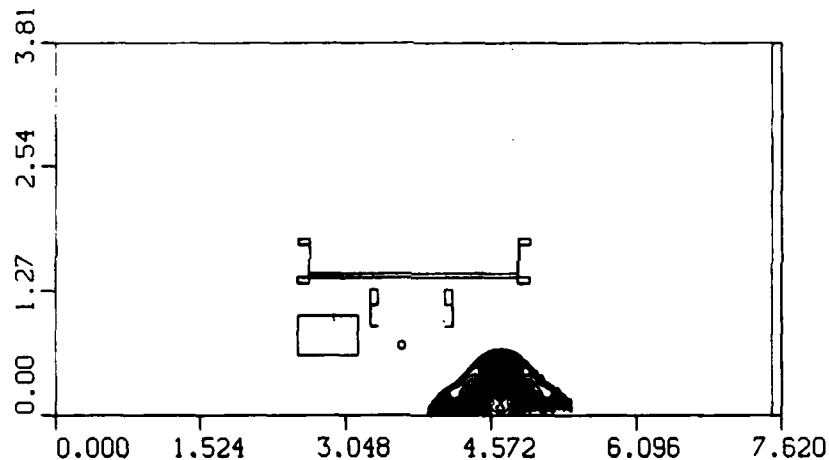


Figure 8. Pressure contours for mine/blast truck cross section interaction.
 $t = 0.0$ msec; $P_{\max} = 158$ MPa; $P_{\min} = 0.099$ MPa.

In Figure 9, the simulation results in the form of pressure contours are shown at the stage when the leading shock wave is impinging on the right corner of the cargo bed. The highest pressure is observed at the point of shock wave reflection and is about 1000 MPa at this stage. Also, very strong reflections are observed at the drive shaft and at the right supporting beam. In Figure 9, very high resolutions of the flow gradients are observable. To assure this resolution, the number of grid points used here was 15,604, or about 6,800 more than in Figure 8. This trend will continue throughout the simulation. Because the flow pattern gets more complex, new shocks and discontinuities appear, requiring more and more computing resources. However, the advantage of the adaptive grid is evident since the new grid points are added only when needed in simulation.

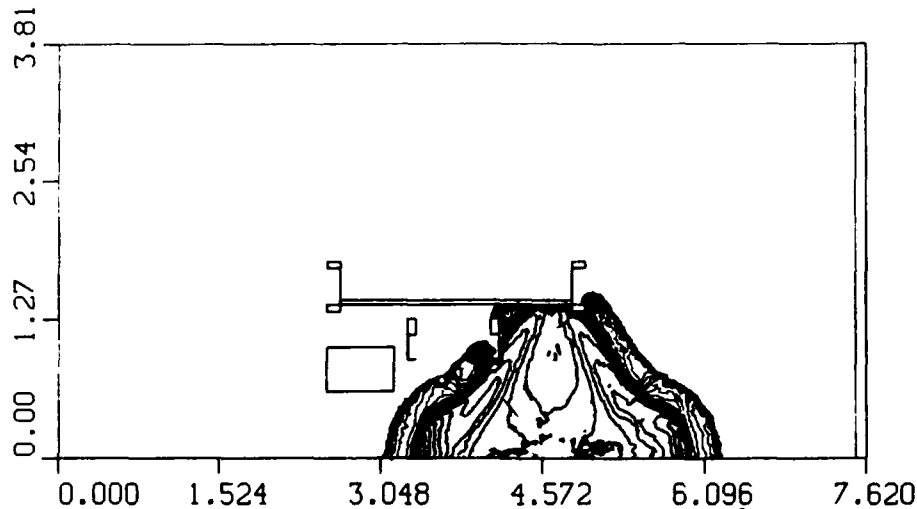


Figure 9. Pressure contours for mine blast/truck cross section interaction.
 $t = 0.147$ msec; $P_{\max} = 953$ MPa; $P_{\min} = 0.7$ MPa.

In Figure 10, the simulation results are shown at $t = 1$ msec. The leading front of the blast wave propagating from the right has reached the cargo bed inner area. At this stage of the simulation, the pressure in the inner volume of the cargo bed is ≈ 0.4 MPa. Highest pressures reaching 25 MPa are observed under the truck in the closed areas that impair venting. The strong shocks at the upper side of the cargo bed are followed by areas of strong rarefactions that lower the pressure to the levels of 0.025 MPa. This combination can produce a severe jolt to the personnel at levels predicted by our simulation, and most probably will be fatal.

In Figure 11, the simulation results are shown at $t = 2.3$ msec. Here, extremely complex flow patterns including tens of shock fronts and other discontinuities are resolved well by the AUGUST code. The collision of the shock fronts from the left and right side of the cargo bed has produced a complex flow that has pressure values up to 0.78 MPa. However, the dominant pressure is about 0.4 MPa. The higher pressures of about 3.6 MPa are observed under the truck in enclosed areas. This high pressure can lead to the truck overturning.

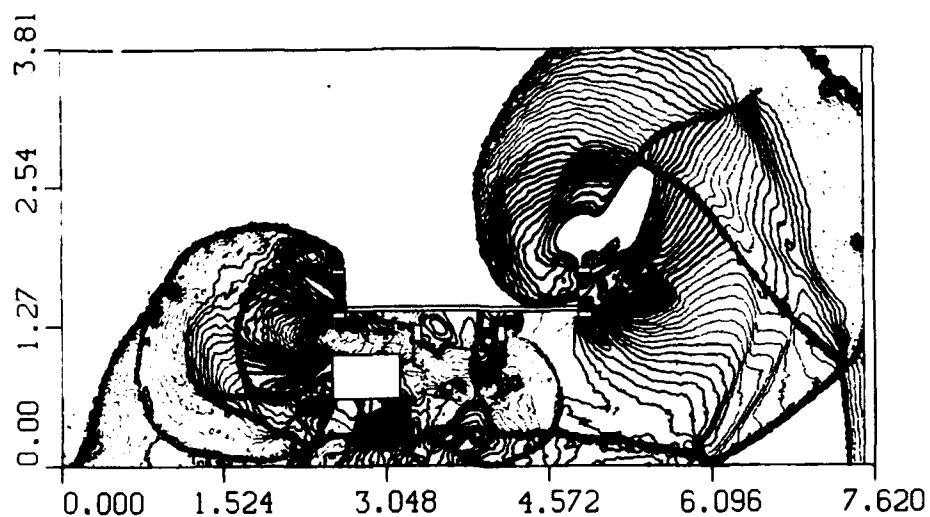


Figure 10. Pressure contours for mine blast/truck cross section interaction.
 $t = 1$ msec; $P_{\max} = 25.8$ MPa; $P_{\min} = 0.025$ MPa.

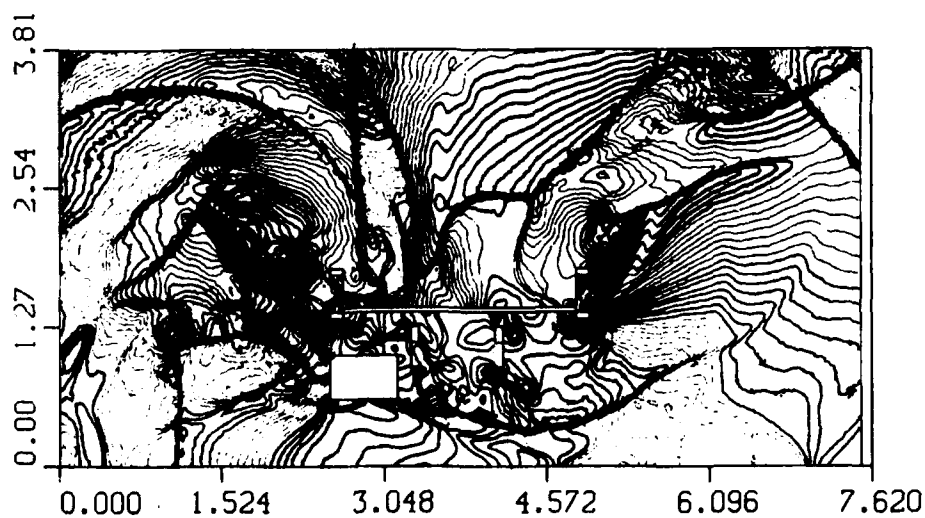


Figure 11. Pressure contours for mine blast/truck cross section interaction.
 $t = 2.32$ msec; $P_{\max} = 3.69$ MPa; $P_{\min} = 0.021$ MPa.

CONCLUSIONS

Increasing the mine protection of the M925A1 cargo truck is an important goal of our program. The geometry modifications that can lead to significant increase in personnel survivability should be minimal and cost effective. This study presented the first results of the truck vulnerability analysis based on detailed numerical simulation of the mine blast-truck interaction. The simulations are performed with a new state of the art code that allow detailed and accurate description of the complex flows over complex geometries. The new AUGUST code is based on adaptive unstructured triangular grids. According to this technique the computational grids dynamically evolve during the simulations, assuring a maximal resolution in the regions of the flow with shock waves and contact discontinuities. The code was validated and comparison of the experimental and simulation results testify to its high accuracy. Using the developed methodology of analysis a base case of the mine blast-truck interaction was simulated. Computed results point to very high pressures that will develop in the area of an unprotected cargo bed. The presented work is a first step in a long range program that will include analysis of the structural response as a result of blast load, and comprehensive analysis of various blast hardening methods.

REFERENCES

1. Mavriplis, D. and A. Jameson, *Multigrid Solution of the Two-Dimensional Euler Equations on Unstructured Triangular Meshes*, AIAA-87-0353, 1987.
2. Baum, J.D. and R. Löhner, *Numerical Simulation of Shock-Elevated Box Interaction Using an Adaptive Finite Element Shock Capturing Scheme*, AIAA Preprint 89-0653, Presented at the AIAA 27TH Aerospace Sciences Meeting, Jan. 8-12, 1989, Reno, NV.
3. Jameson, A., T.J. Baker, and N.P. Weatherill, *Calculations of Inviscid Transonic Flow Over a Complete Aircraft*, AIAA 24th Aerospace Sciences Meeting, Reno, NV, AIAA Paper 86-0103, January 1986.
4. van Leer, B., *Towards the Ultimate Conservative Difference Scheme, V.A. Second Order Sequel to Godunov's Method*, J. Comp. Phys. v. **32**, 101-136 (1979).
5. Eidelman, S., P. Collela, and R.P. Shreeve, *Application of the Godunov Method and Its Second Order Extension to Cascade Flow Modelling*, AIAA Journal, v. **22**, 10 (1984).

BISHOP, SERGI, EIDELMAN, LOTTATI

6. Lottati, I., S. Eidelman, and A. Drobot, *A Fast Unstructured Grid Second Order Godunov Solver (FUGGS)*, 28th Aerospace Sciences Meeting, AIAA-90-0699, Reno, NV, 1990.

7. Lottati, I., S. Eidelman, and A. Drobot, *Solution of Euler's Equations on Adaptive Grids Using a Fast Unstructured Grid Second Order Godunov Solver*, Proceeding of the Free Lagrange Conference, Jackson Lake, WY, June 1990.

8. Collela, P., and P. Woodward, *The Piecewise Parabolic Method (PPM) for Gasdynamic Simulations*, J. Comp. Phys. 54, 174-201 (1984).

9. Glass, I.I. and D.L. Zhang, *Interferometric Investigation of the Diffraction of Planar Shock Waves Over a Half-Diamond Cylinder In Air*, UTIAS Report No. 322, March 1988.

Trapping Characteristics of Radiation-Generated Charge in Silicon-on-Insulator Buried Oxides*

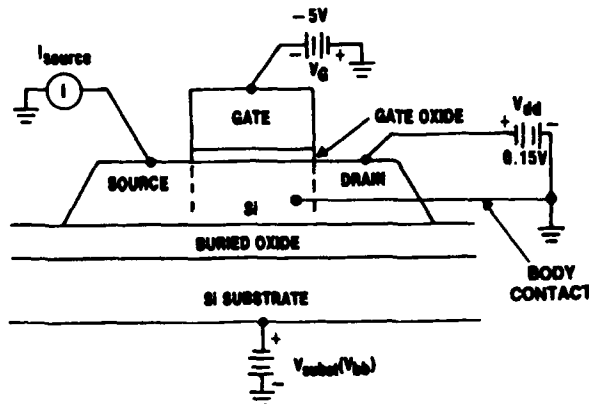
H. Edwin Boesch, Jr., Christine A. Pennise, Thomas L. Taylor
Harry Diamond Laboratories
2800 Powder Mill Road
Adelphi, MD 20783-1197

1. Introduction

Silicon-on-insulator (SOI) technologies show great promise for very high-speed, low-power microelectronic circuits. They also show superior immunity to upset in strategic Army defensive systems in high-dose-rate and single heavy-ion ionizing radiation environments. For these reasons, SOI technologies are undergoing intensive development both in the United States and abroad. Both 64 and 256 K static random access memories (SRAMs) have been produced, and 1 Mbit SRAMs are under development. However, SOI devices are vulnerable to total-dose radiation damage—threshold voltage shifts and backchannel leakage—caused by radiation-generated charges trapped in the buried insulator layer.

Figure 1 shows schematically the configuration of a generic SOI metal-oxide-semiconductor field-effect transistor (MOSFET). It consists of a thin island (100 to 500 nm) of active silicon containing the FET itself (source, drain, and channel region

Figure 1. Schematic of an SOI MOSFET biased to measure its back-channel (buried oxide) $I_s(V_p)$ characteristics.



*This work was sponsored by the Army 6.1 program and the Defense Nuclear Agency.

with overlying insulating silicon dioxide (SiO_2) oxide layer and polysilicon control gate) on top of the buried oxide (BOX) insulating layer* (100 to 1000 nm) that isolates the device from the Si substrate. The SOI structure (silicon-oxide-silicon) may be generated by any of a variety of techniques [1]. The technique that is presently the leading contender for advanced metal-oxide-semiconductor (MOS) SOI circuits is the separation by implantation of oxygen (SIMOX) process [1]. In this process, an SiO_2 layer is produced beneath a surface Si layer by the implantation of a high density (typically, near $3 \times 10^{18}/\text{cm}^2$) of oxygen ions at a depth of several microns into the starting Si wafer. When this operation is performed at moderate temperature and the Si wafer is subsequently heat-treated (annealed) at high temperature (near 1200°C), the result is a thin layer of near-stoichiometric SiO_2 below a thin layer of single-crystal Si of reasonable quality.

It is the use of the BOX layer to isolate active regions from each other that gives SOI technologies their advantage over standard bulk silicon microelectronic technologies. In the latter, devices are *not* isolated electrically by dielectrics but by large-area reverse-biased p-n junctions. These junctions have large capacitance, which reduces circuit speed; require substantial area, which reduces circuit density; and can become forward-biased, which can cause circuit failure by latchup. In addition, they are the source of large photocurrents in high-dose radiation environments (dose-rate upset) and under exposure to high-energy single ions (single-event upset (SEU)). The use of SiO_2 dielectrics for isolation significantly reduces stray capacitances, allows reductions in interdevice dimensions, greatly reduces both high-dose-rate and single-ion photocurrents, and eliminates latchup.

Unfortunately, SiO_2 insulators can trap radiation-generated charges, and these charges (typically, positive charge (holes)), near active regions in the devices, cause changes in device electrical characteristics. For charges trapped in the BOX, the most probable effect is turn-on of a parasitic source-drain leakage channel in the MOSFET Si layer near the interface with the BOX. In SOI MOSFETs using fully depleted Si, the device threshold voltage is also directly affected by charge in the BOX.

In the work discussed here, we have examined the charge-trapping characteristics of the BOX in SIMOX SOI material using two general measurement techniques: (1) x-ray-induced photocurrents, and (2) fast time-resolved current-voltage ($I(V)$) following pulsed radiation exposure. Results of both experiments indicated that BOX produced by the SIMOX process strongly traps holes throughout the oxide bulk [2,3]. In addition, the pulsed measurements revealed strong transient trapping of radiation-generated electrons; most of the electrons were detrapped by a thermally activated process and removed from the BOX within 1 s at room temperature

*In some usages, "SOI" includes silicon devices on insulators other than SiO_2 (e.g., sapphire).

[3]. As discussed in sections 3 and 4, the strong bulk hole trapping and substantial electron trapping of the SIMOX BOX contrasts strongly with the relatively much reduced and primarily interfacial trapping of only the holes that is observed in more conventional high-quality SiO_2 layers grown on the surface of Si at high temperature in an oxygen ambient ("thermal" oxides). As a result of these observations, it has been recognized that the techniques applied successfully in the past to the radiation hardening of the thermal oxide layers of microelectronic devices are not likely to succeed for the SIMOX BOX, and therefore unconventional hardening techniques will be required.

2. Samples and Experimental Techniques

The SIMOX SOI test structures consisted of SOI MOSFETs (fig. 1) supplied by vendors and SOI MOS capacitors supplied by vendors or fabricated by the Harry Diamond Laboratories Semiconductor Electronics Material Technology Center (SEMTC) on SIMOX substrates supplied by vendors. The SIMOX MOSFETs used in the time-resolved $I(V)$ measurements were supplied by Texas Instruments Corporation (TI) and were all taken from a single wafer with a BOX thickness of 420 nm. The capacitor samples used in the photoconduction current experiment were supplied by Ibis Corporation and TI. The thermally grown, radiation-hardened gate-oxide sample was fabricated by Hughes Aircraft Corporation (HAC). TI and Ibis supplied the wafers only, and electrode deposition and packaging were done at the HDL SEMTC. The Ibis BOX samples were nominally 20- or 30-mil diameter dot capacitors with an Al electrode and a BOX thickness of approximately 260 nm. The top Si layer was chemically etched away on these samples, and the Al electrode was deposited directly on the SiO_2 layer. The TI capacitor samples were nominally 18×18 mil square capacitors with a square Al electrode placed on the top Si layer and an oxide thickness of about 432 nm. These samples also had guard rings around the top electrode to eliminate any possible surface leakage paths. The HAC sample was a nominal 25-mil-diameter dot Al-gate capacitor with an oxide thickness of about 39 nm.

3. X-Ray-Induced Photocurrent Measurements

3.1 Background and Theory

The work presented here addresses the issues of the charge transport and trapping properties of the BOX layer. A commercially available 10-keV x-ray source was used to measure the motion of radiation-generated charge using MOS BOX capacitors in a photoconduction current experiment described by Benedetto and Boesch [4]. Here we have used the photocurrent technique as a method to study the trap-

ping characteristics of SIMOX BOX materials. The same 10-keV x-ray radiation source was also used to obtain information about total-dose effects in the BOX by recording pre-and postirradiation capacitance-voltage (C-V) curves and measuring the mid-gap voltage shifts, ΔV_{mg} . The conduction current measurements yield information on the magnitude of the charge *moving* through the oxide layer; however, it is insensitive to the *sign* of the moving charge. Measurements of total-dose-induced ΔV_{mg} provide information on the moment of the net charge *trapped* in the oxide. The results of both the photocurrent measurements and the C-V measurements can additionally be used in theoretical models which provide information about the magnitude and approximate distribution of the radiation-generated charges in the oxide layer.

When an MOS device, under bias, is exposed to ionizing radiation, electron-hole pairs are created in the oxide layer. Depending on the electric field across the oxide, some fraction of the charge carriers escape initial recombination. The generation of charge in the oxide depends on the local electric field, E_{ox} , and the energy, E , of the ionizing radiation. The radiation-generated free electron-hole pairs (the charge carriers) give rise to a conduction current in the oxide layer. If we assume *no* bulk trapping or trapping *only* at the interfaces, then this current is given by

$$I_{co} = Ad[K_g(E)f_y(E_{ox},E)][R_{de}(E,d)\dot{D}_o(E)] , \quad (1)$$

where A is the area of the MOS capacitor (in cm^2), d is the oxide thickness (in cm), $K_g(E)$ the charge generation constant (in C/cm^3 per rad(SiO_2)), $f_y(E_{ox},E)$ the fractional charge yield, R_{de} the dose enhancement factor, and $\dot{D}_o(E)$ the bulk oxide dose rate (in rad(SiO_2)/s)[4]. The fractional yield, $f_y(E_{ox},E)$ is the field-dependent fraction of radiation-generated charge escaping recombination and is derived empirically by Dozier et al [5]. For 10-keV x-rays $K_g^{x\text{-ray}} = 1.38 \times 10^{-6} \text{ C}/\text{cm}^3 \text{ rad}(\text{SiO}_2)$, and R_{de} was determined from existing data for silicon on oxide [4].

The results reported here have been obtained using the photoconduction current technique which measures the current generated in the oxide and is indicative of the charge *moving* in the oxide. It is anticipated that the SIMOX BOX material will contain charge traps. The current that is actually measured by the experiment will be affected by charge traps distributed through the bulk oxide. (The photocurrent measurement is insensitive to interfacial trapping.) Charge traps in the bulk oxide will reduce the distance the charges travel in the oxide which, in turn, reduces the current measured by the experiment. Therefore, due to trapping, the actual photocurrent measured, I_{cm} , in the BOX will be some fraction of the theoretical maximum current given in equation (1). This fraction, I_{cm}/I_{co} , contains information about the amount of charge moving in the BOX, and therefore information about the amount of trapping in the bulk oxide.

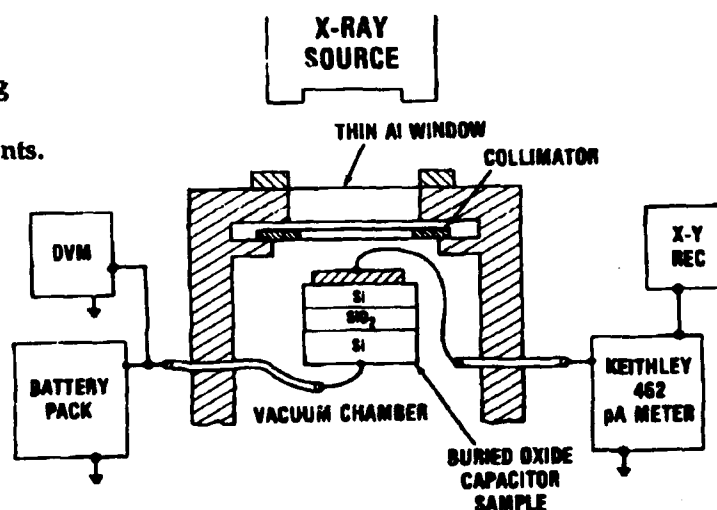
3.2 Experimental Details

All photoconduction current measurements reported in this paper were taken at room temperature using an ARACOR model 4100 10-keV x-ray machine. The samples were placed in a vacuum chamber to eliminate air ionization currents, and a bias was applied to the sample substrate. The guard rings were grounded. A lead collimator was placed above the sample to insure that only the active area of the sample was irradiated, thereby minimizing any stray currents. Upon irradiation, the x-ray-generated current from the buried oxide was measured between the top electrode and ground using a Keithley model 462 picoammeter. Irradiation and measurement time was on the order of a second, which at room temperature makes it impossible to see fast electron trapping and detrapping. A schematic of the experimental setup used to perform these measurements is shown in figure 2. This figure depicts a cross-sectional area of a single MOS capacitor typically used in the photocurrent experiments. The upper superficial Si layer acts as the top electrode, and the lower Si substrate is the bottom electrode.

All measurements were taken with the x-ray machine operating at an accelerating potential of 60 kV. The total dose was kept low to avoid perturbation of the oxide electric field by space charge buildup. Samples were irradiated under oxide fields from about 0.1 MV/cm up to fields where they began to break down (fields of about 3, 4, and 5 MV/cm for TI, Ibis, and HAC samples, respectively).

The same 10-keV x-ray source was used to irradiate the samples in order to measure ΔV_{mg} in the BOX capacitor. The mid-gap voltage shift was found by taking pre- and postirradiation C-V measurements. C-V measurements were made by applying a voltage ramp to the sample substrate and ground and monitoring the sample capacitance with a Boonton capacitance meter connected to the top electrode and ground.

Figure 2. Schematic diagram showing experimental apparatus used to take photoconduction current measurements.



3.3 Results and Discussion

Figure 3 shows a typical set of raw data obtained from the photocurrent measurements. In this figure the x-ray-generated conduction current is plotted versus the oxide field for four BOX samples. As a standard of comparison for the charge motion in the BOX samples, the photocurrent in an HAC radiation-hardened gate oxide capacitor is also shown, since these samples are known to exhibit no measurable bulk trapping. As expected, in each case the conduction current increases with increasing oxide field. Note that figure 3 plots the total conduction currents as measured by the experiment, which are dependent on sample area and oxide thickness, as well as on parameters pertaining to the radiation source, including charge yield, dose rate, and dose enhancement factor (recall eq (1)).

In figure 4 we plot the measured photocurrent I_{cm} normalized to the theoretical maximum current I_{co} as a function of oxide field E_{ox} . As expected, this quantity for the radiation-hardened gate-oxide capacitor is about unity for fields above 0.3 MV/cm. (The slightly higher values at lower fields arise from uncertainties in the measurements and in the empirical formula for f_y as given elsewhere [2].) This is because *both* charges move through the oxide and are either collected at the electrode or are trapped at the interface. Note that for each of the BOX samples, I_{cm}/I_{co} is about the same, despite the differences in processing, capacitor area, oxide thickness, and dose rate. Of even greater importance, observe that for any given field the measured current in the BOX is about half the theoretical maximum value.

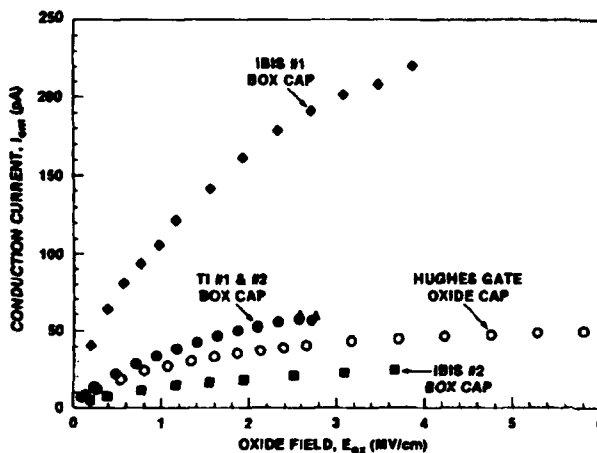


Figure 3. Total measured radiation-induced photoconduction current, I_{cm} , as a function of oxide field.

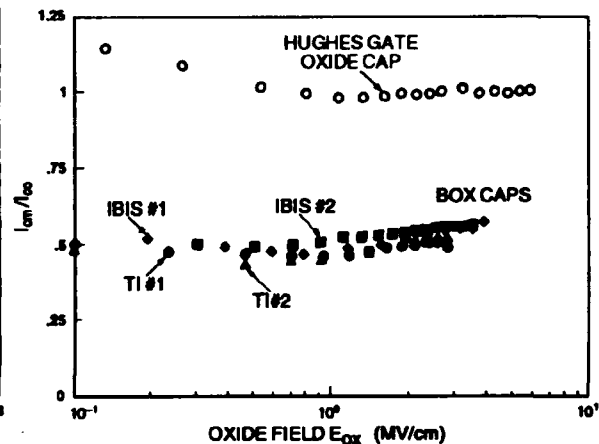


Figure 4. I_{cm}/I_{co} (measured conduction current normalized to the theoretical maximum current) as a function of oxide field.

As noted earlier, if the BOX were behaving like a radiation-hardened oxide and if all the charge were to move through the oxide, the measured photocurrent in the BOX would be the same as that in the radiation-hard samples: i.e., we would measure $I_{cm}/I_{co} = 1$. On the other hand, if one charge is completely trapped at its point of origin while the other moves completely out of the oxide, the result would be $I_{cm}/I_{co} = 0.5$. As shown in figure 4, we actually measure $I_{cm}/I_{co} = 0.5 \pm 0.06$ at all fields for all four BOX samples. In the simplest possible scenario, these results clearly suggest that most likely we have one freely moving and one strongly trapped carrier.

As we have also noted, the photoconduction current measurements provide information on the magnitude but not the sign of the moving charge. In order to determine which carrier is predominately moving and which is predominately trapped, we turn to the C-V measurements that were used to measure ΔV_{mg} . Total dose C-V measurements were performed on two identical TI samples, one irradiated under positive and the other under negative bias, corresponding to an oxide field of ± 1 MV/cm. Pre- and postirradiation C-V curves are shown in figure 5 for samples irradiated to a total dose of 3.24 krad(SiO_2). Mid-gap voltage shifts were measured as $\Delta V_{mg+} = -6.4$ V and ΔV_{mg-} irradiation. The fact that ΔV_{mg} is negative and of fairly similar magnitude for both positive and negative bias is indicative of a net positive charge contained primarily within the bulk of the oxide. The relatively slight difference in the magnitude of ΔV_{mg} in the two cases can be attributed to the hole distribution being skewed toward one interface or the other (depending on bias) and possibly some electron trapping. This implies, as assumed, that most holes are being trapped in the oxide while most electrons are moving out of the oxide. The fact that I_{cm}/I_{co} is so close to 0.5 lends weight to the conclusion that, in fact, the holes are trapped almost *in place* and the electrons move completely through the oxide.

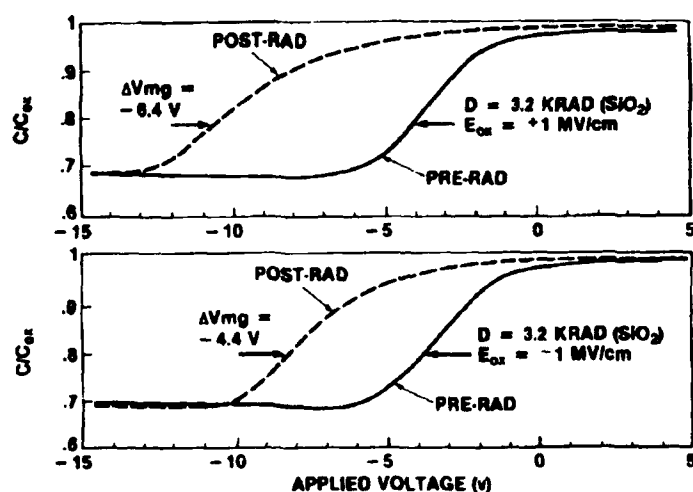


Figure 5. Pre- and post-irradiation capacitance-voltage curves showing the mid-gap voltage shifts for samples irradiated under ± 1 MV/cm.

4. Time-Resolved $I(V)$ Measurements

We measured the time dependence of trapping, detrapping, and transport of radiation-generated charge in the BOX samples using fast $I(V)$ measurements on the MOSFET samples following pulsed irradiation. The experiments were conducted at the Armed Forces Radiobiology Research Institute (AFRRI) electron linear accelerator facility (LINAC). The samples were exposed to single 4- μ s pulses of 13-MeV electrons in an evacuable sample holder which had provisions for temperature control over the range from 77 to 450 K and for active and passive electron-beam dosimetry (thin-foil Cu calorimeter, thermoluminescent dosimeters).

A block diagram of the experimental arrangement appears in figure 6. Since we were interested in charge effects in the BOX rather than the gate oxide of the FETs, the *back-channel* (BOX-top Si interface) rather than the normal gate-controlled FET characteristics were measured (fig. 1). The gate-controlled channel was biased off (gate voltage, $V_{g'} = -5$ V). A drain bias, V_{dd} , of 0.15 V was applied to operate the FETs in the linear $I(V)$ region, and a selected DC *back-gate* voltage, V_{bb} , was applied to the substrate and, unless otherwise indicated, maintained during and after irradiation. We recorded source current, I_s , versus V_{bb} characteristic curves for the test FETs before irradiation and then at logarithmically spaced times from 0.2 ms to 800 s after the radiation pulse by interrupting the DC back-gate bias, applying 150- μ s V_{bb} voltage ramps, and recording both V_{bb} and I_s during the ramps with a high-speed waveform digitizer [6]. Following the measurements, we determined back-gate threshold voltages, V_{th} , from the linear-region $I_s(V_{bb})$ plots by extrapolation of the maximum-slope tangent to the source current axis. We then obtained back-gate threshold voltage *shifts*, $\Delta V_{th}(t)$, as a function of time after irradiation by subtracting pre-irradiation V_{th} from the postirradiation results.

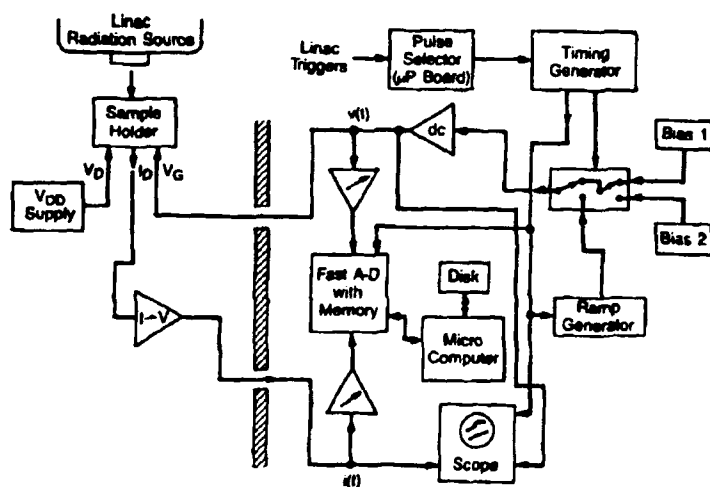


Figure 6. Block diagram of experiment to measure time-dependent charge effects in MOSFETs following short pulse irradiation.

5. Results and Discussion

Figure 7 shows representative pre- and postirradiation $I_s(V_{bb})$ curves for a typical BOX FET irradiated to 9.4 krad(SiO_2) at 295 K with a back-channel bias of -20 V. The postirradiation curves are shifted negatively with respect to the pre-irradiation curve and show an increasingly negative shift with time, indicating the presence of an increasing *net* positive charge in the BOX resulting from the irradiation. The net positive charge is not surprising, since thermally grown SiO_2 insulators typically trap holes but do not show significant trapping of electrons. This result is also consistent with the photocurrent results given in the last section, which indicate strong bulk trapping of holes alone. The *increasing* negative shift is unusual, however, since the negative shift is usually observed to decrease with time as the holes are gradually removed from the oxide [7].

Figure 8 shows $\Delta V_{tb}(t)$ results extracted from data similar to that in figure 7 for samples irradiated with $V_{bb} = +20$ V (fig. 8a) and $V_{bb} = -20$ V (fig. 8b) at temperatures from 137 to 402 K. In all cases, ΔV_{tb} is negative, and in most cases, ΔV_{tb} becomes more negative with time. This behavior under both positive and negative bias cannot be explained by the loss of holes or their movement in the BOX. To explain the observed negative $\Delta V_{tb}(t)$, we must conclude that a *substantial density of radiation-generated negative charge* (electrons) is also present and moving out of or being detrapped from the BOX at the temperatures and over the time ranges of our measurements. In this respect, the SIMOX BOX is very different from conventional thermally grown oxides: in those materials, the radiation-generated free electrons are swept out of the oxide within picoseconds after their generation, and no significant trapping or detrapping of electrons is observed.

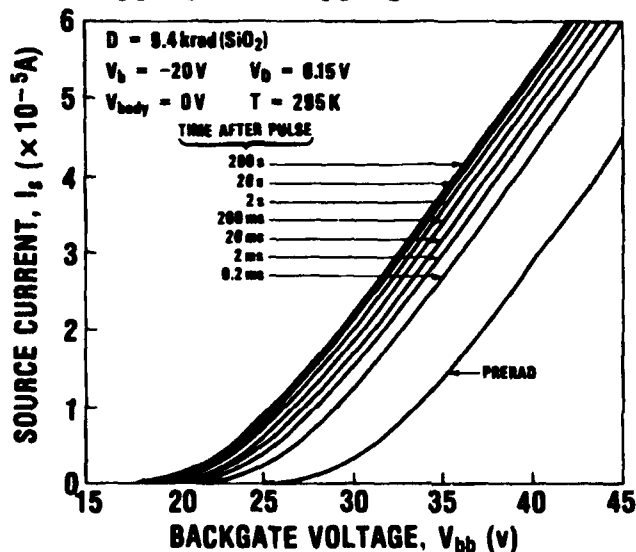


Figure 7. I_s versus V_{bb} characteristics of SIMOX FET following irradiation with $V_{bb} = -20$ V.

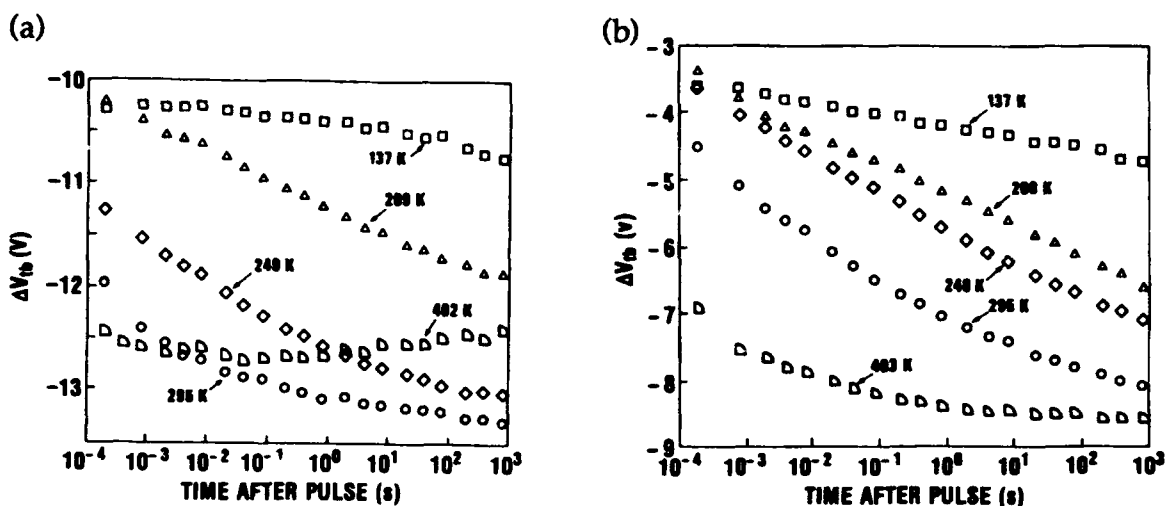


Figure 8. ΔV_{th} as a function of time after radiation pulse exposure for SIMOX FETs irradiated at various temperatures: (a) $V_{bb} = +20$ V during and following irradiation and (b) $V_{bb} = -20$ V during and following irradiation.

5.1 Early and Late ΔV_{th}

Figure 8 also shows that $\Delta V_{th}(t)$ depends strongly upon temperature and the sign of the bias, V_{bb} . Examination of the time and temperature dependences suggests that whatever the process that results in increased negative ΔV_{th} with time, it is barely beginning at 0.2 ms at 137 K, is in various stages from 209 to 295 K, and is completed by 800 s at 402 K. The initial ΔV_{th} 's before this process takes place appear to be about -10.2 and -3.2 V for V_{bb} equal to 20 and -20 V, respectively. At late times and higher temperatures, the ΔV_{th} 's approach about -13.5 and -8.5 V for 20 and -20 V, respectively. (For the 20-V, 402 K case, the data show a turnaround that probably arises from a different process.) Presumably, the early-time ΔV_{th} values result from the net oxide charge present after any very fast charge transport and trapping processes are complete and before the process responsible for the increasing negative shift takes place. Since ΔV_{th} is proportional to the moment of the charge distribution in the BOX, the differences in initial shifts under positive and negative V_{bb} indicate that the net positive charge distribution is asymmetric and is strongly skewed toward the top Si under positive V_{bb} and toward the substrate Si under negative V_{bb} . Therefore, some process must have taken place during or following the initial uniform generation of charges in the BOX to cause this asymmetric distribution of charge related to the electric field applied across the BOX. The most likely process is early charge motion followed by trapping.

5.2 Bulk Charge Trapping Model

If we assume that both electron and hole traps are distributed uniformly through the BOX and we ignore space-charge effects and hole-electron recombination at traps, then we may derive expressions for the expected shifts, ΔV_+ and ΔV_- , under positive and negative bias, respectively:

$$\Delta V_+ = -3.82 \times 10^6 f_y D \{ S_h [d \exp(-d/S_h) - S_h(1 - \exp(-d/S_h))] + S_e [d - S_e(1 - \exp(-d/S_e))] \} , \quad (2)$$

$$\Delta V_- = 3.82 \times 10^6 f_y D \{ S_e [d \exp(-d/S_e) - S_e(1 - \exp(-d/S_e))] + S_h [d - S_h(1 - \exp(-d/S_h))] \} , \quad (3)$$

where S_h and S_e are the mean free paths for capture of holes and electrons respectively, and D is the oxide dose. After substituting our initial ΔV_{tb} values, -10.2 and -3.2 V, for ΔV_+ and ΔV_- , and appropriate values for the constants, these equations can be solved simultaneously and yield values of 25 and 202 nm for S_h and S_e . The low value for S_h implies very strong trapping of the holes—on the average they move less than 6 percent of the BOX thickness before being trapped and are captured essentially uniformly through the oxide. On the other hand, the relatively large value of S_e indicates that the electrons may move through about half the BOX thickness before being trapped. They are therefore trapped in a distribution skewed toward the positive electrode. We conclude that immediately following short-pulse irradiation of the BOX, we have a skewed distribution of trapped electrons superimposed on a nearly uniform distribution of trapped holes. Therefore the trapping characteristics of the SIMOX BOX contrast strongly with the observed characteristics of conventional thermal oxides. In addition to the electron trapping, SIMOX shows extremely strong trapping of holes in the bulk of the BOX, whereas thermal oxides typically show substantial hole trapping only near the SiO_2 -Si interfaces [7].

5.3 Electron Loss Processes

We now examine the physical processes responsible for the postirradiation time-dependent ΔV_{tb} —the loss of electrons from the oxide. The candidate mechanisms for this slow loss or recovery process are (1) a slow transport of the electrons through and out of the oxide, similar to that for holes in thermal oxides, or (2) a thermal detrapping process in which the electrons are thermally excited from traps and then drift rapidly out of the oxide.

In high-quality thermal oxides, the holes move by means of a relatively slow time-dispersive polaron hopping process that is probably an intrinsic property of

amorphous SiO_2 [8,7]; in thicker, less radiation-hard materials, multiple trapping/detrapping processes (trap hopping) may also be involved [9,10]. The stretchout of the recovery over many decades in time in the transport results from the variable distance from one hopping site to another—the holes move essentially by tunneling from site to site, and the time scale for the tunneling varies exponentially with distance. However, electrons have been found to be quite mobile and transport very rapidly in thermally grown SiO_2 ; hopping transport of electrons has not been observed. A characteristic of the dispersive hole transport in thermal SiO_2 is that the shape of the $\Delta V_t(\log t)$ recovery curves (similar to fig. 8) is largely independent of temperature; that is, the curves are simply shifted along the log time axis by changes in temperature. In addition, a single activation energy applies to the process from its early stages (recovery just beginning) to its late stages (recovery nearly complete).

In early work, Simons and Hughes [11,12] proposed a thermal detrapping model for "short-term annealing" or negative ΔV_t recovery in MOSFETs by detrapping of holes. This model assumes that the charge carriers are captured in traps with a spread in energy, and the ΔV_t recovery over many decades in time results from the exponential dependence of carrier release time on trap depth. Following release from the traps, the carriers are assumed to drift rapidly out of the oxide with no retrapping. A thermal detrapping process of this type results in recovery curves that vary in shape with temperature and show a substantial increase in activation energy from the early to late recovery as the recovery proceeds from the shallow to the deeper traps.

5.4 Activation Energy Analysis

Data from figure 8b ($V_{bb} = -20$ V) were analyzed for activation energy of the electron loss process using the method of crosscuts. Assuming that removal of 100 percent of the electrons results in the observed change in ΔV_{tb} from its early value of -3.2 V to its late value of -9.0 V, we drew crosscuts through the data at ΔV_{tb} values corresponding to fractions of 25-, 50-, and 75-percent recovery. The intersections with the data yielded the time to reach the various recovery fractions at the various temperatures. These recovery times are plotted in figure 9 on a log scale as a function of inverse temperature (Arrhenius plot). Reasonable straight lines could be fit through the recovery time points at each recovery fraction; the slopes of these curves then yielded the apparent activation energy, E_a , as shown in the figure. The large increase in E_a (from 0.29 eV at 25-percent recovery to 0.92 eV at 75-percent recovery) contrasts strongly with the shift of only a few hundredths of an electron volt observed over this recovery and temperature range for polaron hole transport in gate oxides [10]. This suggests that electron transport by an analogous process is prob-

ably not the electron recovery mechanism. (Dispersive electron transport resulting from a thermally activated multiple-trap-hopping process is not ruled out, however.) Instead, the wide range in E_d indicates that a spread in electron trap depths is probably responsible for the extended recovery process—i.e., thermally activated detrapping.

We then analyzed the data in figure 8b using the formalism presented by Simons and Hughes [11,12] for hole detrapping. Assuming a single detrapping process (no multiple trapping/detrapping), a spatially uniform distribution of electron traps, and certain other conditions, the distribution in energy, $P(E_{fp})$, of the trapped electrons is given by

$$P(E_{fp}) = (2\epsilon/kTqd) [d(\Delta V_{tb})/d(\ln t)]_{E_{fp}}, \quad (4)$$

where

$$E_{fp} = kT \ln(st), \quad t > 1/s, \quad (5)$$

and ϵ is the BOX dielectric constant, q is the electronic charge, and s is an attempt-to-escape frequency for the electrons in the traps. Following Simons and Hughes [12], s is assumed to be constant at $1 \times 10^{11} \text{ s}^{-1}$. Since s enters as part of the argument of the log, E_{fp} is not a strong function of s , and a change in s simply shifts the scale for E_{fp} . The value of $d(\Delta V_{tb})/d(\ln t)$ was determined graphically from the data in the figure for one-decade time increments at each temperature; the corresponding values for E_{fp} were also calculated. The results are plotted in figure 10 as charge density, P , as a function of energy, E_{fp} . The points obtained from the various temperatures are in fair

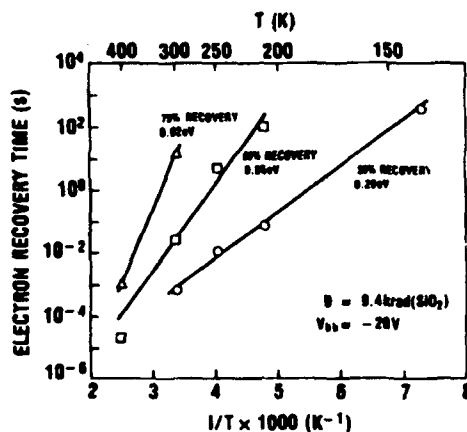


Figure 9. Electron recovery time as a function in inverse temperature for SIMOX FETs irradiated with $V_{bb} = -20 \text{ V}$ for 25-, 50-, and 75-percent recovery.

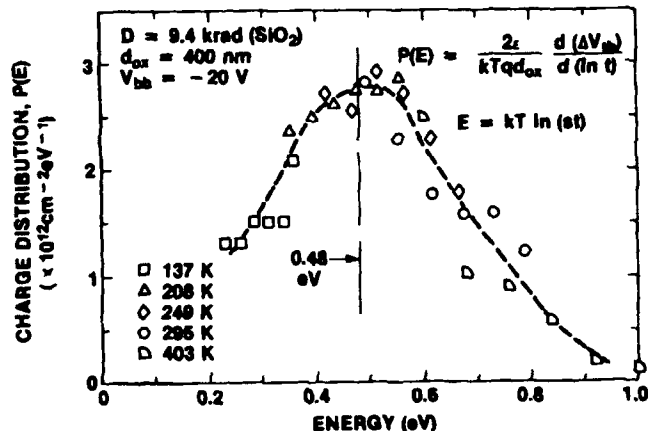


Figure 10. Distribution of radiation-generated electrons in SIMOX FETs irradiated with $V_{bb} = -20 \text{ V}$ as a function of thermal detrapping energy.

agreement, and the overall result is a roughly symmetrical broad peak centered on about 0.48 eV. The absolute value of $P(E_{fp})$ should not be taken too seriously, since it is based on the assumption of spatially uniform electron trapping, and we know the actual spatial distribution is skewed (toward the back channel, in this case). As noted above, the absolute value of the trap energies may also be in error through the uncertainty in s . The assumption of single detrapping with no retrapping is also suspect, given that S_e is less than the BOX thickness. Despite these qualifications, the reasonable results obtained here strongly suggest that thermal detrapping is the key process in the electron recovery.

6. Conclusions

X-ray photoconduction and time-resolved $I(V)$ measurements have been performed on representative SIMOX buried oxide SOI structures in order to determine the nature of the transport and trapping of radiation-generated charge in the BOX layer.

The photoconduction current measurements were made on MOS capacitors using fairly simple and straightforward experimental techniques with commercially available equipment. The photocurrent measurements together with C-V shift measurements indicate that the holes are essentially trapped in place and that most of the electrons move completely through the oxide.

The time-resolved $I(V)$ measurements confirm the strong trapping of holes in the SIMOX BOX. In addition, these measurements indicated substantial initial trapping of electrons followed by rapid detrapping of most of these electrons in less than a second at room temperature. The electrons are evidently thermally detrapped from a broad distribution of shallow traps.

Overall, the results clearly indicate that the SIMOX BOX differs substantially in radiation response characteristics from more conventional thermally grown oxides. As a consequence, the radiation hardening techniques used for conventional oxides may not be applicable, and new hardening techniques may be required.

Acknowledgments

The authors would like to thank Bohdan Dobriansky and Judith McCullen of the HDL SEMTC for sample preparation; Barry McLean, Tim Oldham, and Jim McGarrity for useful discussions; and the AFRRI LINAC staff for their support and electrons.

References

1. T. D. Stanley, "The State-of-the-Art in SOI Technology," *IEEE Trans. Nucl. Sci.* NS-35, 1346 (December 1988).
2. H. E. Boesch, Jr., T. L. Taylor, L. R. Hite, and W. E. Bailey, "Time-Dependent Hole and Electron Trapping Effects in SIMOX Buried Oxides," *IEEE Trans. Nucl. Sci.* NS-37, 1982 (December 1990).
3. C. A. Pennise and H. E. Boesch, Jr., "Determination of the Charge-Trapping Characteristics of Buried Oxides Using a 10-keV x-Ray Source," *IEEE Trans. Nucl. Sci.* NS-37, 1990 (December 1990).
4. J. M. Benedetto and H. E. Boesch, Jr., "The Relationship Between ^{60}Co and 10-keV X-Ray Damage in MOS Devices," *IEEE Trans. Nucl. Sci.* NS-33, 1318-1325 (December 1986).
5. C. M. Dozier, D. M. Fleetwood, D. B. Brown, and P. S. Winokur, "An Evaluation of Low-Energy X-Rays and Cobalt-60 Irradiations of MOS Transistors," *IEEE Trans. Nucl. Sci.* NS-34, 1535-1539 (December 1987).
6. J. M. Benedetto, H. E. Boesch, Jr., F. B. McLean, and J. P. Mize, "Hole Removal in Thin-Gate MOSFETs by Tunneling," *IEEE Trans. Nucl. Sci.* NS-32, 3916-3920 (December 1985).
7. F. B. McLean, H. E. Boesch, Jr., and T. R. Oldham, Chapter 3 of *Ionizing Radiation Effects in MOS Devices and Circuits*, T. P. Ma and P. V. Dressendorfer, eds., Wiley (1989).
8. F. B. McLean, H. E. Boesch, Jr., and J. M. McGarrity, "Dispersive Hole Transport in SiO_2 ," Harry Diamond Laboratories, HDL-TR-2117 (1987).
9. J. R. Srour, S. Othmer, O. L. Curtis, Jr., and J. Y. Chiu, "Radiation Induced Charge Transport and Charge Buildup in SiO_2 Films at Low Temperature," *IEEE Trans. Nucl. Sci.* NS-23, 1513-1519 (December 1976).
10. O. L. Curtis, Jr., and J. R. Srour, "The Multiple-Trapping Model and Hole Transport in SiO_2 ," *J. Appl. Phys.* 48, 3819 (1977).
11. M. Simons and H. L. Hughes, "Short-Term Charge Annealing in Electron-Irradiated Silicon Dioxide," *IEEE Trans. Nucl. Sci.* NS-18(6), 106-113 (December 1971).
12. M. Simons and H. L. Hughes, "Determining the Energy Distribution of Pulse-Radiation-Induced Charge in M-OS Structures from Rapid Annealing Measurements," *IEEE Trans. Nucl. Sci.* NS-19(6), 282-290 (December 1972).

BRAUN

Advanced Electronic Detonator Modules for Live-Fire Missile Testing (U)

* Captain Christopher G. Braun, Ph.D.
U.S. Army Electronics Technology and Devices Laboratory
(LABCOM) SLCET-ML
Fort Monmouth, New Jersey 07703-5601

Introduction

To accomplish its mission, the Army is continually developing and fielding new weapons and battlefield systems. A vital part of this process is testing under simulated battlefield conditions to assess both the capability of U.S. weapons as well as vulnerability of our systems to both enemy and U.S. weapons. The Live-Fire and Vulnerability Division of the Combat Systems Test Activity (CSTA) located at Aberdeen Proving Ground is one of several U.S. centers which conduct live-fire and vulnerability testing. As mandated by Congress before fielding a new weapon system or as the need arises, program managers will task CSTA or other centers to conduct an evaluation of a particular weapon or assess the vulnerability of a system. Typically, production or near production ordnance is used and a large number of shots are fired. To exactly control the velocity, angle of incidence, position and response upon impact, as well as to limit the associated costs with a large test range, these tests often restrict the operation of the weapon. In particular for the work at CSTA, missiles under test are strapped to a sled on a guided rail and accelerated by external rockets.

One side effect of this approach of missile testing is that it becomes nearly impossible to use the "safe & arming" (S&A) devices which are built into each missile to prevent accidental triggering (Figure 1). This results from the complexity of the safe & arming devices which are much like mechanical watches with gears, safety stops and inhibits. The function of this S&A device is to render the munition safe from triggering from accidental events including shock, vibration, electrical noise or heating possibly caused by circumstances ranging from dropping the munition during loading to operating near a radio transmitter. S&A devices are designed not to trigger until after several events have occurred indicating that a valid launch has taken place. It also ensures that the munition cannot trigger until sufficiently far away from the soldiers launching the missile such that they cannot be harmed by detonation.

BRAUN

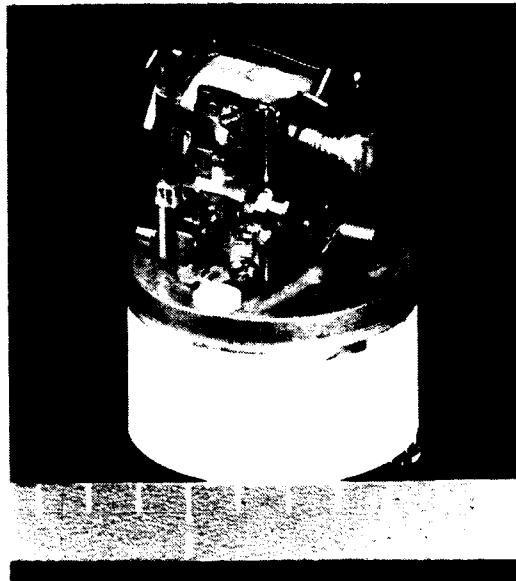


Figure 1. A safe & arming (S&A) device from a TOW missile system. A hole in the center cylinder holds the low energy detonator (not shown) 90° away from the warhead until after launch. It is then mechanically rotated in line and ignites the warhead on impact.

The function of S&A devices is yet more complicated due to the nature of triggering detonators and explosive warheads. Detonators come in many sizes, shapes and "sensitivities". In general, there are low energy detonators, which require only a few microjoules to trigger, and high energy detonators, which require a large fraction of a joule to trigger. The more sensitive a detonator is the easier it is to produce a valid trigger and ignite the explosive charge. However, it is also then intrinsically more sensitive to all forms of accidental triggering. One example is the M48 low energy detonator which is used in production TOW missiles.¹ It requires 12.5 microjoules to trigger which is normally supplied by a low voltage signal generated from closing a crush switch in the missile nose cone by impact. But this same trigger energy can come from a static electric discharge, for instance by walking across a carpet, or from dropping the device or from heating. CSTA has used these low energy detonators on some test programs. However, extraordinary safety measures were required, such as placing a humidity chamber around the missile to prevent a static charge accumulation.² Even so, with the number of personnel involved and the extensive handling of the missile, safety concerns regarding low energy detonators can never be fully resolved for test programs.

These issues force CSTA remove or bypass the safe & arming devices from the missiles they test. Because of the controlled nature of the test situation, the S&A devices in the missile systems will not arm and thus will not trigger the detonator to

BRAUN

ignite the warhead upon impact. Because of the complexity of these S&A devices it is impractical to modify them. Because of safety concerns it is unwise to use a low energy detonator without the safe & arming device to keep it protected from false trigger signals and (in some cases) keep it physically removed from the explosive charge until after launch. Hence, another approach is required to meet the safety and performance constraints of missile test programs.

One of several valid alternatives is to use a high energy detonator. This has the safety advantage of being nearly immune to accidental ignition due to noise, shock, heating, etc. It has the strong disadvantage that it is so insensitive that the triggering circuit requires an electrical pulse of over 300 kW (>600 V, >500 A) with a risetime of 10^9 amps per second. This is greater than 4 orders of magnitude (10^4) more power or energy than is required by the low energy detonator. Consequently, the electronic triggering circuit for the high energy detonator transforms from trivial to a near impossibility in the constraints for on board a missile. However, one of CSTA's major goals is to preserve test realism implying that the internal crush (impact) switch of the missile must be used. This goal constrains the detonator electronics to be mounted on board with the missile.

In 1990 a representative from CSTA presented this challenge to us at the Pulse Power, and asked if we could apply the technology we were developing for the Army in power semiconductors to this need. At that time, we were in the stage of performing device evaluation and characterization. We had no experience in applications with this new technology, especially not for missile systems where many working constraints would be placed upon a design. Over the past two years, this project has grown from a notion that it might be possible, to a vital part of the Improved HELLFIRE missile test program, to the point today where it is an intrinsic piece of the CSTA live-fire testing methodology.³

This paper describes the work that has taken place in developing this new kind of electronic detonator module for use by the missile test community. The road of taking a technology from the white-halled laboratory environment to in-the-field-and-dirty Army use is often tortuous. My observation is that technology does not transfer easily. Further, end users and research laboratories have quite different philosophies: users avoid risk since it imperils their programs whereas laboratories embrace new technologies. This project has survived and succeeded only because of the personal commitment by the believers at CSTA and ETDL, not because it was good or needed nor because of any technology master plan.

BRAUN

Requirements

The goal has been to develop an electronic detonator module for use on board a missile capable of reliably firing a high-energy Exploding Bridge Wire (EBW) detonator. CSTA performed some preliminary work with an SCR-based detonator modules, but they were found to be unreliable.³ Given the cost of each test, including the missile and target, can exceed \$200,000 it is vital from an economic consideration that a detonator module work every time. As in most cases where new methods and technologies were being applied, the requirements evolved as we gained experience in the operating environment.

The requirements for an electronic detonator module include:

- High energy EBWs (RP-80) necessitate a discharge pulse ≥ 300 kW
- Small, lightweight and rugged so as to place inside missile under test
- Highly reliable so as to give high confidence in results and eliminate retesting costs and time of both missile and target
- Completely self-contained for power for ~ 4 sec
- Able to discharge a $0.5 \mu\text{F}$ capacitor charged to 1050 V; $di/dt \geq 1,000 \text{ A}/\mu\text{S}$
- Trigger voltage ≤ 30 volts
- Automatically rendered safe after power down in case of an abort
- Able to monitor module status up to launch to ensure correct operation
- Safe from not triggering due to noise pickup, static discharge, etc.
- An external safety short in parallel to the RP-80 must prevent detonation

Reynolds Industrial Systems Incorporated is the only U.S. manufacturer of exploding bridge wire detonators. The RP-80 EBW detonator is a small (2 cm x 0.5 cm) cylinder with two leads to connect the to the electronic trigger module. These detonators are placed in a cavity in the warhead charge. Applying a sufficiently strong electrical pulse causes the explosive material in the detonator to ignite which then leads to the detonation of the warhead.

RP-80 Exploding Bridge Wire (EBW) detonator:⁴

- 78 mg Pentaerythritol Tetranitrate (PETN) - 1/4 stick TNT
- Bridgewire 1.5 mil gold
- All-fire ≥ 384 kW
- No-fire ≤ 76 kW
- A study showed that discharging a $0.5 \mu\text{F}$ capacitor charged ≥ 823 V will reliably fire the RP-80



BRAUN

Switching Technology

The mission of the Pulse Power Center at the Electronics Technology and Devices Laboratory is to develop and transition to applications critical components for high electrical powers. Mostly, this work focuses on development of high-power switches as they are subject to the highest stresses and are generally the weak link of any power system. ETDL has been a part of five-year, multi-million dollar, coordinated tri-service and industry effort to develop MOS Controlled Thyristors (MCTs) as the new generation of semiconductor power switch.⁵⁻⁷ Consequently, when CSTA presented their requirements we saw an opportunity to begin to apply MCT technology to Army needs.

In general, there are several means to meet the requirements to trigger a EBW detonator. All standard circuits will use a capacitor bank to store energy at ~ 1 kV and a fast, high-power switch to dump this energy into the EBW. During normal operation a 500 A pulse of about 1.2 μ s FWHM is produced with a current rise time of about 1,000 A/ μ s. For a fault or for firing into the safety short, the requirements jump to a 1,000 A pulse in a 1 μ s FWHM pulse with a di/dt of about 3,000 A/ μ s. The constraints upon size, weight, ruggedness and performance begin to eliminate some possible choices for switching technologies.

In the context of high power switching technologies there are three families of switches which could meet the performance requirements: krytrons, Insulated Gate Bipolar Transistors (IGBTs) and MCTs. Nuclear weapon triggers use krytron switches, a normally open plasma device, with EBW detonators for safety and timing purposes. A trigger voltage of 200 volts initiates the plasma on the grid of the krytron and the switch rapidly closes. It has the advantage of all gas switches of high holdoff voltage, fast switching times, ruggedness and high immunity to radiation. Krytrons have the disadvantage of cost per device, a high voltage trigger circuitry, concerns about the susceptibility to shock and of lifetime for applications which are not single shot. IGBTs have emerged as the standard medium power semiconductor switch, capable a few hundreds of amps with a blocking voltage of up to 1500 V. These devices are commercial and readily available. IGBTs have the drawback that they are linear devices and do not latch on (as a thyristor or a plasma switch) which results in a less efficient on-switch and one that would be relatively large for the peak current (1 kA) required in a fault condition. Further, IGBTs are almost exclusively made by Japan and there are serious concerns about dependence upon foreign suppliers for critical military components.

Since a MCT is essentially a power SCR controlled on and off by a dense surface array of MOSFET gates, it has the advantage of thyristors that it is a highly

BRAUN

modulated latching on-device and can handle up to 4 kA in a die of 0.3 cm². Prototype MCTs have been produced with blocking voltages up to 1700 volts for epitaxial devices and up to 3000 volts for diffusion-doped devices. The latter are far from commercialization whereas the epitaxial devices are in limited commercial production. MCTs have the serious drawback that they are still in the development stage, causing device parameters to change, and they are not available on the commercial market. We chose to use MCTs in this application both because they were a good surge turn-on device but also because this application would be an opportunity to begin to transition this DoD-sponsored technology to the field.

The MOS Controlled Thyristor (MCT) is a new high power solid-state switch developed by DoD for advanced power systems:

- MCT is a turn-on and turn-off switch
- Operates over -100 °C to +150 °C range
- Peak surge current up to 4 kA, di/dt >4 kA/us
- Fast switching: 100 ns turn on, 2 μs turn off
- Blocking voltage up to 1700 volts (typical 600-1400 V)
- Very low energy to trigger (~ FET): +10 V off, -10 V on
- Very immune to di/dt or dV/dt accidental triggering



There exist several different MCT device designs which change the ratio of the turn-on FET cells to the turn-off cells. The older MCTs were optimized for turn-off and had a lower ratio of turn-on to turn-off cells. Testing at ETDL showed that a 1/25 (turn-on to turn-off cells) MCT could not withstand the high rate of rise of current and were destroyed during the switching process. However, 1/9 and 1/4 design MCTs are able to handle di/dts of greater than 4,000 A/μs and have been successfully used in detonator modules. Because of the lack of commercial production of MCTs at this time, CSTA placed a contract with Harris Semiconductor for one production lot (1000 to 4000 switches depending upon yield). These switches are optimized turn-on N type 1/4 MCTs with a designed blocking voltage of 1400-1600 volts. We expect that these devices will be more capable than the older P type MCTs as the blocking junction is now a N-doped layer resulting in faster switching speed, better efficiency and increased safe operating area.

MCT Detonator Module Operation

The detonator module is essentially a low inductance capacitor discharge circuit with an off bias to ensure the module does not fire until the crush switch trigger is valid. Shown on Figures 2 and 3 are a photograph of the detonator module and a circuit layout, respectively.

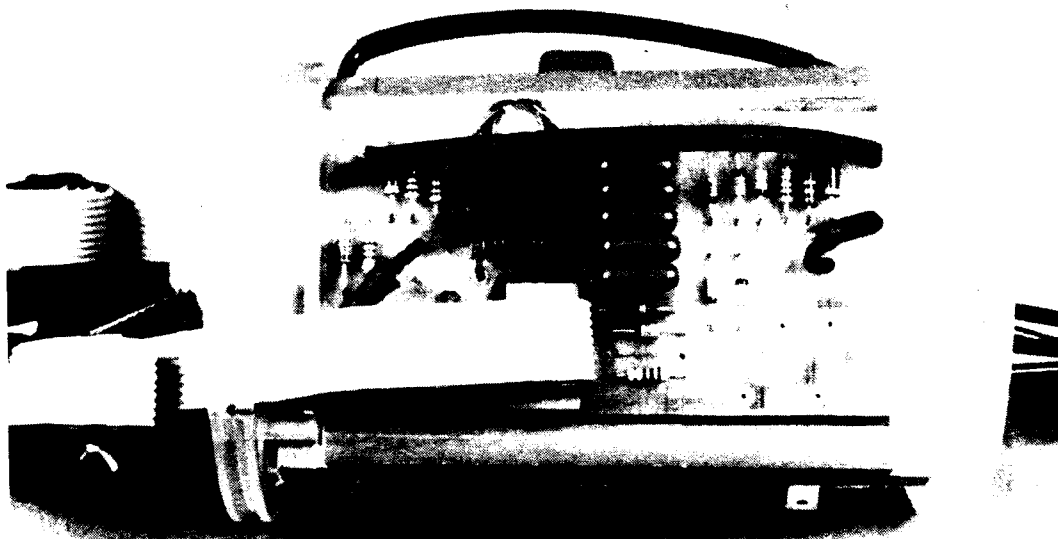


Figure 2. MCT-based electronic detonator module for exploding bridge wire detonators. The MCT high-power switch is the black chip center. The module dimensions are 8 x 4.5 x 2 cm.

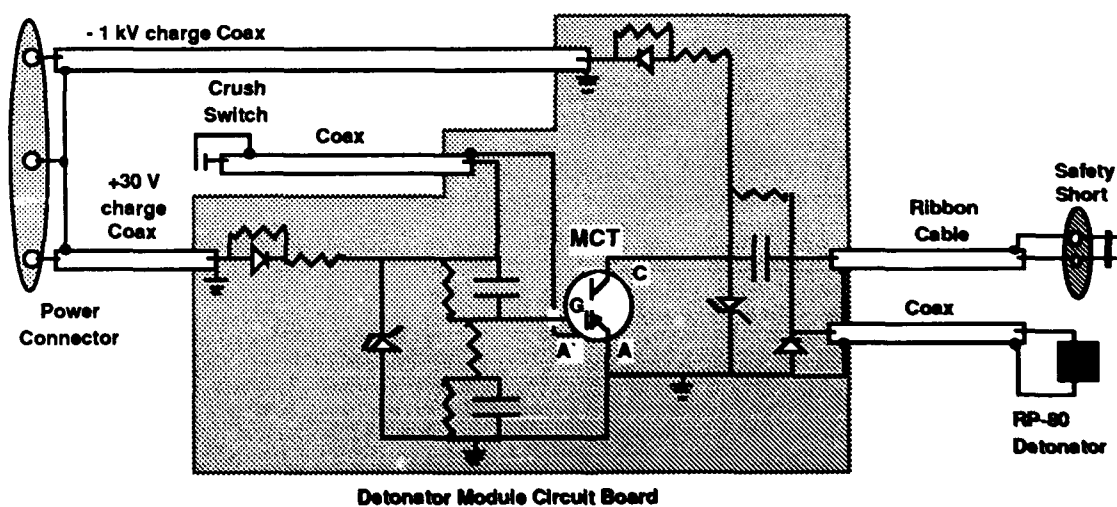


Figure 3. Circuit diagram of the MCT-based electronic trigger module for exploding bridge wire detonators.

BRAUN

In preparation for operation in a missile system the MCT detonator module is placed in a aluminium protective can which is filled with high strength epoxy with only the electrical leads exposed. This can is mounted in the missile body with the input power leads attached to an external connector. The crush switch cable is attached to missile internal crush switch. The safety short cable is attached to an external shorting bar. Just prior to launch the RP-80 detonator is attached to the output cable of the module and then inserted into the main explosive charge. All personnel are cleared from the area to a bombproof shelter.

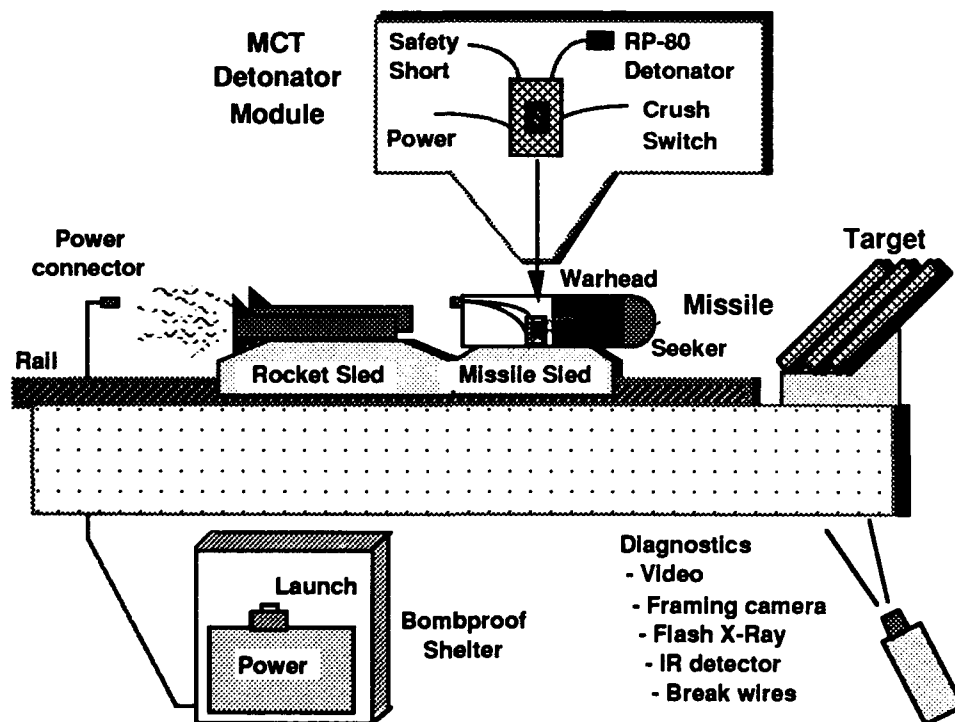


Figure 4. Shown is a schematic of the missile test system at CSTA. The missile is mounted upon a sled on a rail to guide it to the exact target location and accelerated by rockets to the desired velocity. Upon impact the crush switch in the missile closes and triggers the MCT module which outputs a high power pulse igniting the detonator causing detonation of the warhead.

Figure 4 shows a system layout for missile testing with all major components - detonator, missile, rocket sled and rail, target and diagnostics, and bombproof shelter with the power supplies. Figure 5 is part of the missile firing documentation showing an Improved HELLFIRE (IHW) missile on the rail at the target being prepared for launch. Power is applied to the detonator module a few seconds prior

BRAUN

to launch charging the internal gate and pulse store capacitors. In this state the MCT is held by a gate bias voltage in the off mode which continually turns off current (up to 60 amps) - keeping the module insensitive to spurious signals from static discharges, RF pickup or even diagnostic X-ray flashes. If there is an abort prior to launch the detonator module will bleed off the stored energy and render the missile safe to approach. The rocket sled motors are ignited by applying power from a trigger box in the bombproof shelter and the sled quickly accelerates down the rail. As the sled moves from the launch site, the power connector to the missile pulls away and the detonator module is under its own power. Before impact the external safety short is mechanically stripped away by shear plates along the rail.

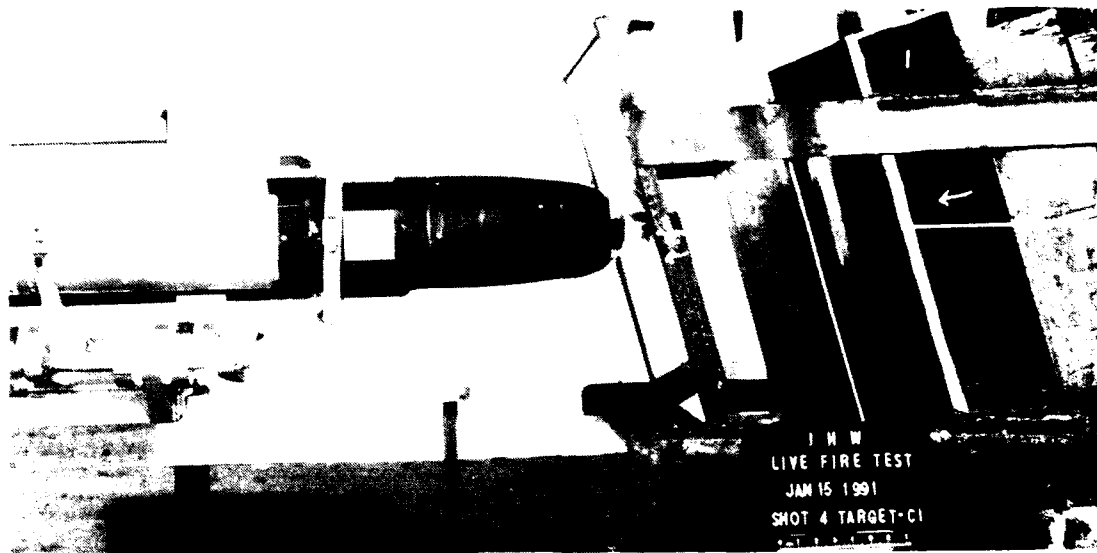


Figure 5. Preparation for one of the Improved HELLFIRE missile firings using a MCT-based electronic detonator module. The first two plates at the target are triggers for diagnostic instrumentation.

Upon impact at the target the missile crush switch changes from open to closed, causing the gate circuitry of the detonator module to produce a turn-on trigger to the gate of the MCT. Within about 100 ns current will begin to flow into the RP-80 detonator. By 600-800 ns the current has reached ~500 amps and the power in the load is over 300 kW (Figure 6) - the fine gold wire in the RP-80 detonator vaporizes and produces a shock wave which causes the surrounding PETN explosive material to begin to burn. In 2-4 μ s, depending upon geometry and burn rate, the explosion will break out of the RP-80 cylindrical enclosure and couple into the main warhead. This starts a chain of explosive burning in the shaped charge warhead resulting in an intense, focused high velocity (5-10,000 m/s) plasma jet forming. The jet then moves forward and penetrates the target (Figure 7).

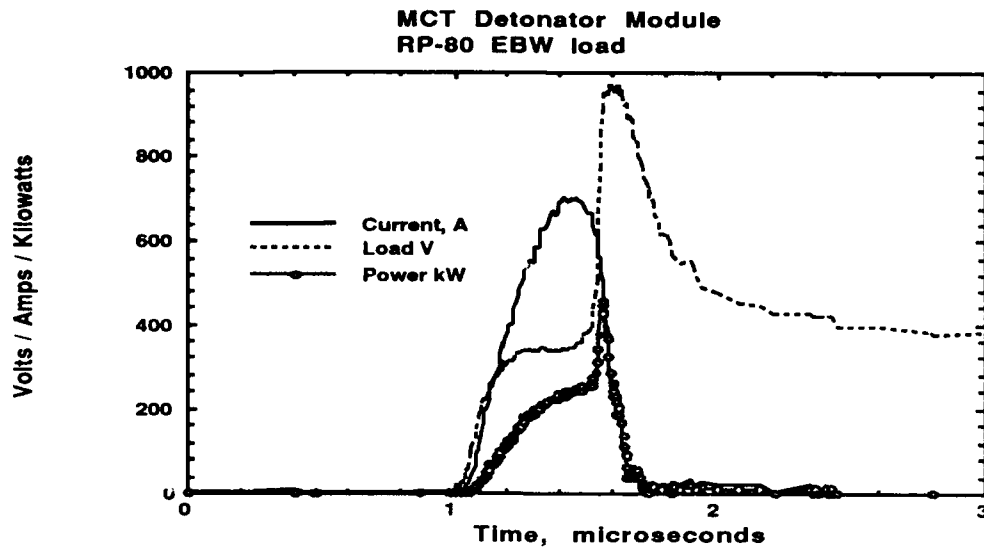


Figure 6. Measurements of the current, voltage and power of a MCT-based detonator firing a RP-80 EBW. The bridge wire vaporizes in the detonator at 0.6 μ s after the start of the current pulse and ignites the PETN explosive in the RP-80.



Figure 7. Detonation of a HELLFIRE missile using an MCT-based electronic detonator.

BRAUN

The performance of this MCT detonator module has been excellent. In comparison to other methods of electronically triggering EBWs for missile shots which have a success rate of 60-80%, the cumulative performance for rail shots is over 92 %. This detonator module has also been successfully used in reactive armor testing where a tank round is fired at reactive armor tiles. These tiles are triggered by one or more MCT detonator modules at appropriate time intervals.

A summary of test results to date:

- | | |
|---|---------|
| • TOW missile live-fire and other practice rail shots | 6 / 6 |
| • HELLFIRE live-fire rail shots | 19 / 21 |
| • Reactive armor testing triggering one or two RP-80s | 30 / 30 |

Tandem Module Development

Many of the newer U.S. missiles have two or more explosive charges to counter threat systems which have added layers of reactive armor. The first charges engage the reactive armor tiles which then allows the main warhead charge to function properly and penetrate the target. An upcoming test of the Optimized HELLFIRE missile (HOMES) requires the use of a tandem electronic detonator module to fire the two explosive charges at preset time intervals. A sequencer module is now under development which provides safety inhibits and control of the timing of multiple events. This sequencer module when combined with two MCT-based detonator modules forms a tandem detonator as shown on Figure 8.

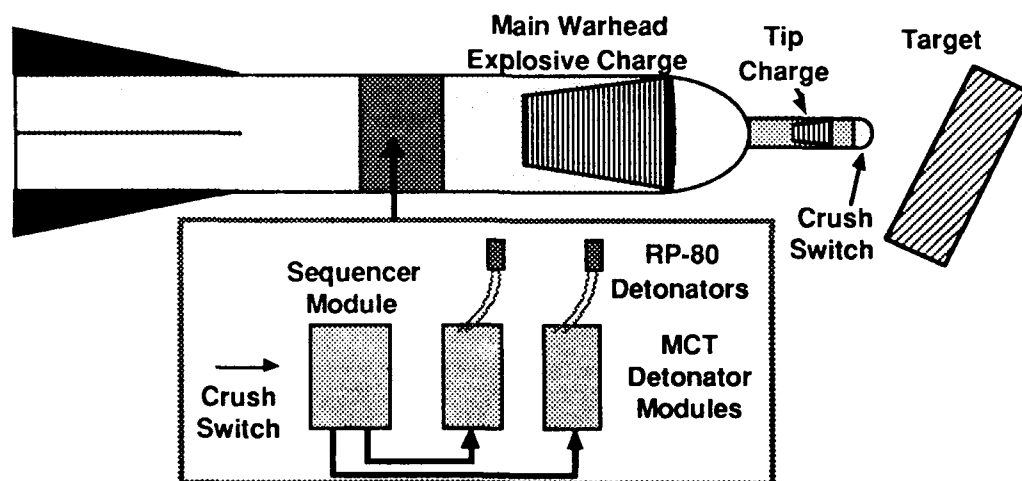


Figure 8. Shown is a tandem detonator module for dual charge missile. A sequencer module provides control of the two MCT detonator modules.

BRAUN

The prototype sequencer module in a tandem configuration under development is shown on Figure 9 (the two MCT detonator modules are hidden behind the sequencer module in the photograph). The complete tandem module is housed in a plexiglass cylinder of dimensions 11 cm in length and 8.5 cm in diameter. Design changes in the near future to the sequencer and detonator modules will permit much better packing and thus smaller system size.

The design approach to the sequencer module was to contain all the logic and timing information in a Erasable Programmable Logic Device (EPLD) leaving only the power, input, clock and output sections constructed in hardware. An EPLD is a relatively new kind of programmable chip which can replace a large number of separate logic chips, limited only by the number of internal macrocells. This approach to the sequencer allows modifications quickly and easily without changing any hardware functions. The internal logic for the EPLD in a sequencer module consists of enable circuitry, a 24 bit counter and three 24 bit comparators. Time intervals for the time from launch to arm the missile and for the delays between outputs T0, T1 and T2 are set by assigning binary values to the comparator inputs in the design for the EPLD. With a 2 MHz clock, the time delays can range from 0 to 4 seconds with a resolution of 0.5 μ s. Each of the outputs T0, T1 and T2 can trigger a MCT detonator module or other event as desired.

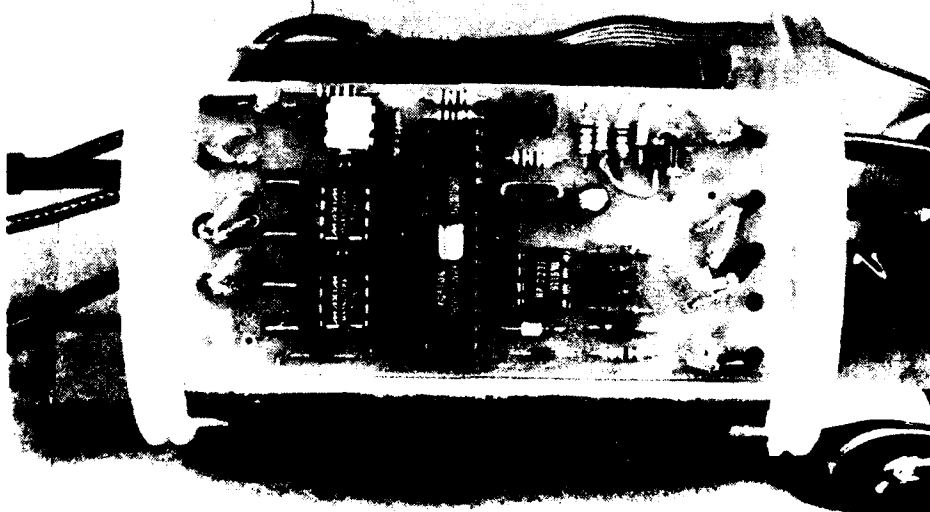


Figure 9. Shown is the sequencer module which provides the safety inhibits and timing of up to three events (detonators). The input section is shown on the right; the output FETs are to the left. The long chip in the center is the EPLD logic device.

BRAUN

Initial testing of the tandem module has shown successful operation, however it has also demonstrated that the electrical noise generated by the detonator modules (dB/dt, RF pickup and ground bumping) can also cause the sequencer to reset. As the detonator modules fire, the fast current rise gives rise to a variety of effects including floating the local ground by tens of volts, generation of a time varying magnetic field and transient reflections through the common grounds due to instantaneous changes in the RP-80 detonator impedance as the explosive ignites. Work is currently underway to resolve the source of the noise and to better isolate the sequencer from it. To prove correct operation for a live-fire missile testing, a TOW IIA missile will be fired with this tandem module. Figure 11 shows a time sequence picture of firing a two explosive charge missile with this tandem electronic detonator module.

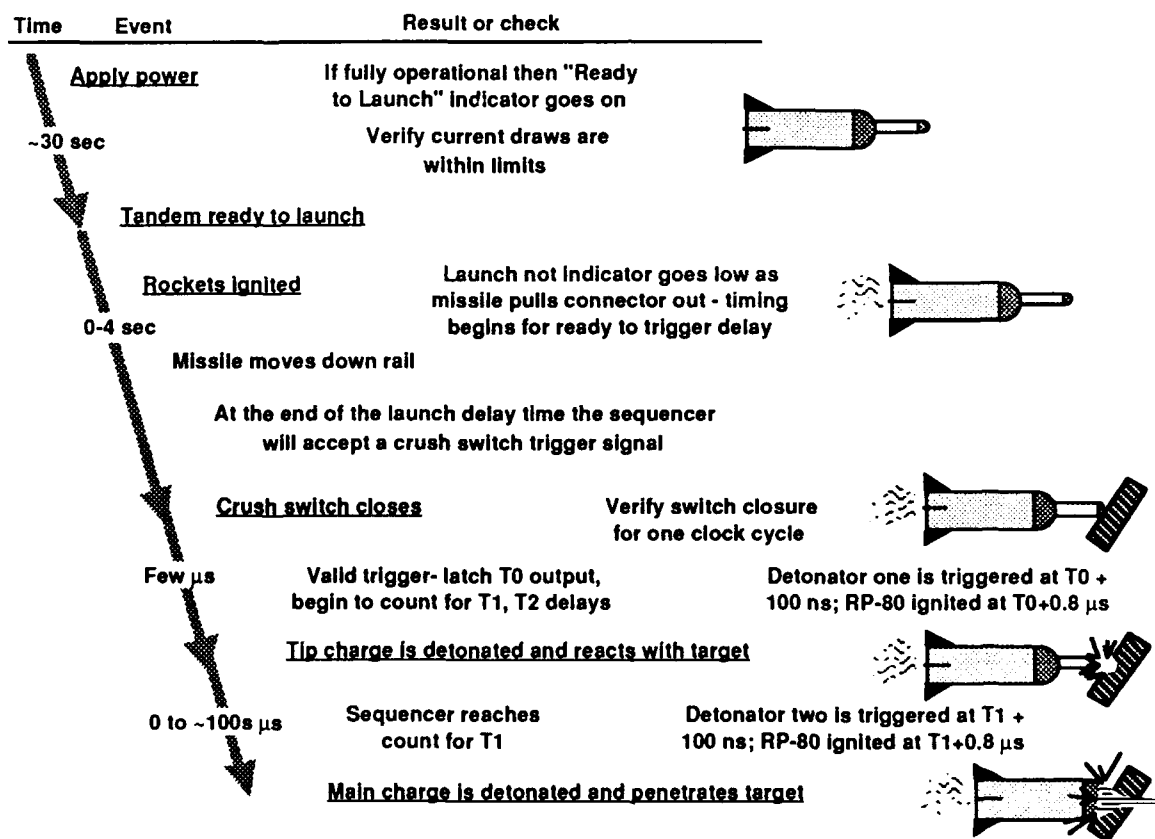


Figure 10. Shown is the time sequence of events for launch and operation of a tandem missile test.

BRAUN

Capabilities of prototype sequencer module:

- Provides 3 output events from 0 to 4 seconds in steps of 0.5 μ s
- Outputs can trigger detonators or other events (i.e., retro-rockets)
- Input and output are buffered and isolated
- All logic / timing in Programmable Logic Device (PLD)
 - Easy software changes to timing
 - Can reprogram functional logic to make major changes without hardware changes
- Programmable lockout time after launch to disable crush switch input
- Completely self powered for 30 seconds without a battery
- Diagnostic output channel gives timing of all events
- Output status line "Ready to Launch" indicates to operator that sequencer module is operating correctly prior to launch

Summary

The progress in the past two years in developing and proving successful operation of the MCT-based detonator module has been tremendous. This project has demonstrated a new application area for high-power electronic, and has taken our switching technology out of the laboratory and put it to good use for the Army.⁸ We have shown that a MCT-based detonator module is capable of surviving on board a missile system and reliably igniting EBW detonators. This module has been used in the improved HELLFIRE test program at CSTA with a success rate of 19/21 for live fire missile shots, and used in static firings for reactive armor testing with a success rate of 30/30. The key technology, an MCT high power switch, is the result of several years of development by ETDL and other DoD research agencies. This switch makes the detonator module more compact, reliable and affordable than could otherwise been possible. Additionally, a sequencer module which provides the safety inhibits and timing for several events is under development. When finished, this and the detonator module will provide a never-before-available means to realistically test multiple charge missiles. Such types of electronic detonators may be incorporated into future production weapon/armor systems to eliminate the mechanical safe & arming devices in current systems and to provide the ability to dynamically respond to target composition.

The Army shall always need to test weapons and systems for lethality and vulnerability to ensure that our forces are well equipped to engage any foe. This project has contributed significantly to the ability of the U.S. test community to test and evaluate many classes of missiles and other explosive devices.

BRAUN

Bibliography

- [1] D. M. Levin, U.S. Army Combat System Test Activity, "Development of a Reliable TOW Missile Warhead Arming Device for use in Live Fire Vulnerability Testing," IDD-87-03, (1987).
- [2] C. T. Davey, Franklin Research Center, "Examination of the Safety Aspects of the M48 Detonator Used with the TOW Missile Tests at Aberdeen Proving Ground," (1987).
- [3] D. M. Levin, U.S. Army Combat System Test Activity, "MCT Warhead Detonation Circuit," Report USACSTA-7134, TECOM Project 7-CO-R90-APO-001, (1991).
- [4] Reynolds Industries Systems Inc., "Exploding Bridgewire Ordnance," Applications notes and technical data sheets.
- [5] R. Pastore, C. G. Braun, M. Weiner, and S. Schneider, "Developmental MOS Controlled Thyristor (MCT) Behavior," Proc. 19th Power Modulator Symposium, (1990).
- [6] R. Pastore, C. G. Braun, M. Weiner, "Characterization of 3000 V MOS Controlled Thyristors," Proc. 8th IEEE Pulsed Power Conference, (1991).
- [7] S.D. Arthur, V. A. Temple, "High Power Switch Final Report," DAAL01-87-C-0719 (1991).
- [8] C. G. Braun, "A MOS-Controlled Thyristor Based Exploding-Bridge-Wire Detonator Module," Patent Office docket 07/766,592, (1991).

Acknowledgements

There are several engineers who have contributed greatly to the success of this project including Jerry Cook at the Combat System Test Activity, Matt Davenport of the Instrumentation Division, and Bill Wright of Signum Electronics. Don Levin, also of the Instrumentation Division, deserves special recognition for his persistence in originating this project and keeping it alive.

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

Catalytic Antibodies as Biological Scavengers for
Organophosphorus Poisons

Dr. A. A. Brimfield*, Dr. D.E. Lenz, Mr. D.M. Maxwell and
Dr. C.A. Broomfield

Biochemical Pharmacology Branch

U.S. Army Medical Research Institute of Chemical Defense
Aberdeen Proving Ground, Maryland 21010-5425

INTRODUCTION Recent research has been reported on the use of injectable protein scavengers that can react with the toxic organophosphorus nerve agents on a one to one molar (stoichiometric) basis and remove them from circulation. When rhesus monkeys were given equine butyrylcholinesterase or fetal bovine serum acetylcholinesterase prior to soman, they were protected against 2 MLD₅₀ (1) or up to 4 MLD₅₀ of soman (2) respectively with no signs of toxicity or post exposure behavioral deficit. Equine butyrylcholinesterase also afforded protection in monkeys challenged with one MLD₅₀ of sarin. These results have been reviewed recently (3).

Although stoichiometric scavengers offered promising results, there were several drawbacks to their use. Each molecule of scavenger reacted with only one molecule of nerve agent, thus requiring the administration of large doses of protein to achieve the protective effect. If protection against a recurring exposure to a chemical threat was desired, stoichiometric scavengers would have to be administered repeatedly which is an additional disadvantage.

We have now begun work that capitalizes on the striking effectiveness of the biological scavenger approach to protection by turning our attention to biological scavengers that possess catalytic activity. Molecules of that type would render a toxic organophosphorus poison non-toxic by catalyzing its hydrolysis. Since they would repeatedly hydrolyze new substrate molecules, catalytic scavengers could provide protection with smaller doses of exogenous protein. Additionally, they would remain active for the length of their circulatory life and provide continuing protection against agent exposures. While a recent report on the utility of using an enzymatic scavenger to protect mice (4) has been supportive of the concept of catalytic biological scavengers, the human gene for this enzyme has not yet been expressed.

With the advent of this new concept, we felt that an examination of the potential of catalytic monoclonal antibodies as biological scavengers was well justified. The model on which our antibody was based was the hydrolysis of fluorophosphate and fluorophosphonate esters which is known to proceed via a trigonal bipyramid transition state (5). This report summarizes our efforts to date to produce a monoclonal catalytic antibody capable of hydrolyzing the

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

highly toxic nerve agent soman.

MATERIALS AND METHODS

Chemicals. Our transition state analog (TSA) was 4-amino-3,3-dimethyl-2-butoxy-2-methoxy-2-methyl-1,3-dioxo (4,6-di-*tert*-butyl-2-benzo) phosphole synthesized under contract by R.M. Moriarty and his associates in the Department of Chemistry at the University of Illinois at Chicago (6) (Figure 1). The phosphole was conjugated to bovine serum albumin (BSA) through an amide linkage to yield an immunogen with an epitope density of 38 moles/mole. An additional conjugate, 4-aminophenyl pinacolyl methylphosphonofluoridate diazotized to keyhole limpet hemocyanin (KLH-GD), was used in hapten affinity purification, and was obtained from Ash-Stevens, Detroit, MI. Soman was obtained from the Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD. ^{14}C -soman (pinacolyl[^{14}C]methylphosphonofluoridate; sp. act. 54.5 mCi/mmol) with a purity greater than 90% by thin-layer chromatography was purchased from Los Alamos National Laboratory (Los Alamos, NM). Immunochemicals were purchased from Sigma Chemical Co. (St. Louis, MO). Other chemicals were of analytical grade and were commercially obtained.

Animals. Female BALB/c mice, 20-22 g, were obtained from Charles River Laboratories (Wilmington, MA). They were housed in facilities approved by the American Association for the Accreditation of Laboratory Animal Care. They received Purina Lab Chow (Purina Mills Inc, Richmond, IN) and tap water *ad libitum*.

Immunization. Mice were immunized by ip injection with the transition state analog conjugated to BSA carrier protein. They received 50 ug (protein) of conjugate in saline emulsified in Freund's complete adjuvant in a final volume of 0.1 mL. The animals were rested for two months and were boosted with 50 ug of the same conjugate by the ip route in Freund's incomplete adjuvant. Three days prior to fusion, mice received a final boost of 50 ug per animal, ip, in 0.85%

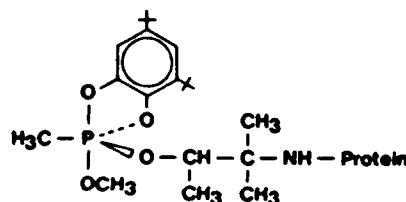


Figure 1. Structure of the transition state analog.

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

saline. The state of immunization was monitored using Enzyme Linked Immunosorbant Assay (ELISA) on Immulon 2[®] microtiter plates (Dynatech Laboratories, Chantilly, VA) coated with a TSA conjugate to human serum albumin, complimentary to the BSA conjugate used for immunization. Detection employed the alkaline phosphatase system.

Fusion. Cell fusion for hybridoma generation was polyethylene glycol mediated (1300-1600 Mol. Wt.) with HAT selection according to a standard protocol (7). The non-secreting SP2/O-Ag-14 myeloma (8) employed as a fusion partner was a gift from Howard Meyer of our Institute.

Identifying Catalytic Monoclonals. Testing of the nascent hybridoma cultures to identify those producing antibody with the desired catalytic activity was carried out in a visually scored colorimetric screening assay based on the release of H⁺ during soman hydrolysis and the resultant change in pH. It is a system developed in our laboratory by Little, et al. (9) modified for use with Multiscreen[®] microtiter plates (Millipore Corp., Bedford, MA) which have nitrocellulose membrane in the bottom of the wells in place of the plastic. Briefly, 200 uL of tissue culture fluid was sterilely removed from the hybridoma culture plate and pipetted onto the membrane prewetted with saline. The sample was drawn through the membrane using a vacuum manifold, also supplied by Millipore, and the membrane was rinsed five times by filling the well with saline and drawing it through the membrane. Antibodies in the culture fluid are bound to the membrane and potential sources of buffering action are washed away. A solution of 2×10^{-4} M soman in 0.85% saline containing 20 uL/mL of methyl red (1 mg/mL in ethanol) was added and the plate was incubated at room temperature for several hours. A positive control was included consisting of a dilute solution of the organophosphorus acid anhydride hydrolase cloned from *P. diminuta* (4) in fresh tissue culture medium. Fresh tissue culture medium alone served as a negative control. Positive wells were amber to rosy red in color and negatives were lemon yellow.

Cloning. Cultures from positive wells were expanded and cloned by limit dilution (10). Clones were reevaluated with the methyl red assay and ELISA. The most active clone, designated IIA12 ID10, was mass cultured for immunoglobulin production.

Immunoglobulin Purification. Culture fluid was diluted 1:4 in Tris buffered saline (TBS) filtered through a 0.45 micron filter (Nalge Co., Baltimore, MD), and affinity purified by a modification of the method of Ey, et al. (11) with KLH-GD replacing protein A as the immobilized ligand. Eluted immunoglobulin was dialyzed overnight against 4 L of TBS, and concentrated against high molecular weight polyethylene glycol. Protein

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

determination was by O.D. 280 using an extinction coefficient of 1.4/mg. For purposes of kinetic assay, some antibody preparations were exhaustively dialyzed against 0.85% saline.

Immunoglobulin Typing. Antibody class and subclass determinations were done by Ouchterlony double diffusion assay using a commercial kit available from The Binding Site (San Diego, CA).

Kinetic Analysis. The method used for our kinetic assays was described by Broomfield, et al. (12). Briefly, enzyme and radiolabelled phosphate or phosphonate ester substrate are combined in a suitable low volume container (we used microtiter plates) at the desired concentrations and incubated. Incubation is terminated by pipetting the contents of each reaction vessel into a tube containing 1 mL of toluene and 1 mL of distilled water. The tube is vortexed, the phases allowed to separate, and the phases are drawn off by Pasteur pipette and added to scintillation vials for quantification of the radioactivity in the aqueous and organic phases. In the assay presented here, four substrate concentrations were run in triplicate both with antibody (0.1 mL, 39 ug/mL in 0.85% saline) and without. The counts from the non-antibody containing wells were subtracted from the antibody catalyzed counts to correct for "non-enzymatic" hydrolysis. The data were analyzed by standard enzyme kinetic techniques using double reciprocal plots.

RESULTS

The immunization regime was successful. Sera from the immunized mice titrated out to greater than a 1: 100,000 dilution when they were tested on a microtiter plate coated with the HSA conjugate of the transition state analog (Figure 2). No estimate of the fusion efficiency was possible because of bacterial contamination of the fusion which was brought under control with antibiotics.

At two weeks post fusion, culture supernatant from the remaining fusion wells was tested for catalytic activity. Several produced positive results, but the cultures growing in two of the wells, designated II A3 and II A12, stood out. These were cloned immediately and supernatant collection from expanded cultures was begun to initiate affinity purification. Additional assays for catalysis indicated that activity increased with purification on both the KLH-GD hapten and on protein A Sepharose, a protein isolated from *Staphylococcus aureus* that binds immunoglobulins independent of their specificity. The protein A indicated that the source of our activity was an antibody. Material originating with the II A12 culture, and purified on each of these immunoabsorbents was tested for mouse immunoglobulin type and subclass. The results of this test were positive for the presence of mouse immunoglobulin, kappa light chain character (kappa is the most prominent species among the light chains produced by mice), and IgG2a. The outcome

relative to type and subclass was the same for either method of purification.

The kinetic experiments were carried out with KLH-GD affinity purified material from one of the clones of the II A12 culture designated IIA12 ID10. The basic steps in the experiment are outlined in Materials and Methods. Additional details are found in the caption to Figure 3. Counts done prior to the start of the analysis showed that the substrate soman was 7.4% hydrolyzed. In the toluene-water system used to separate soman from the hydrolyzed product, pinacolyl methylphosphonic acid, the hydrolysis product and its associated radioactivity are extracted into the aqueous phase. The results for the aqueous counts in the experimental and control samples were corrected to reflect this preexisting contamination.

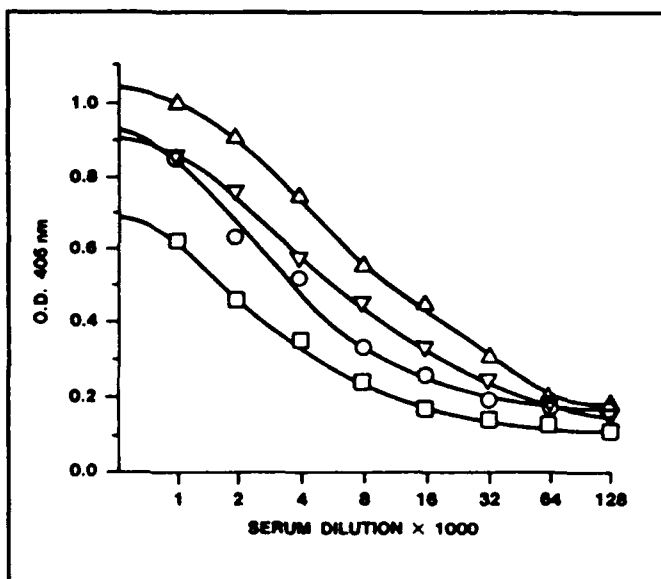


Figure 2. Mouse serum titers after immunization with the BSA-transition state analog.

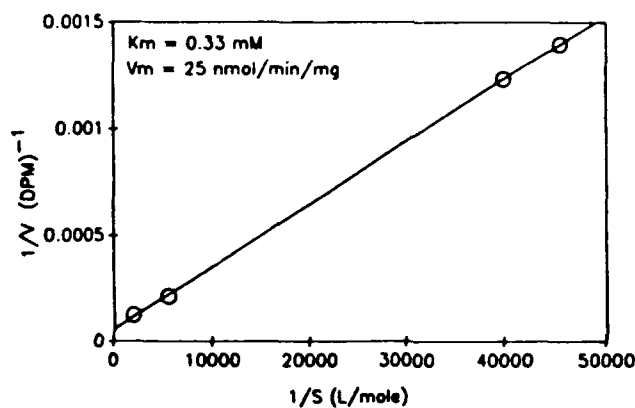


Figure 3. Double reciprocal plot of C^{14} -soman hydrolysis by IIA12 ID10. Samples were run in triplicate. Other conditions were: temperature 26°C; 3.9 μ g of antibody protein per sample; molar soman concentrations, 5×10^{-6} , 1.8×10^{-6} , 2.5×10^{-6} , 2.2×10^{-6} , 120 min incubations.

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

V_{\max} equalled $25 \text{ nmol min}^{-1} \text{ mg}^{-1}$ protein, K_m was $3.3 \times 10^{-4} \text{ M}$, and k_{cat} , defined as V_{\max} divided by the total antibody concentration, was 4 min^{-1} .

DISCUSSION

Given our interest in proteins with catalytic activity, we felt that the recent development of catalytic monoclonal antibodies presented us with a unique opportunity. Immunoglobulins (antibodies) represent a totally different class of physiologically benign proteins that lend themselves to use *in vivo*. Antibodies in crude sera or in purified form have been used therapeutically for snake bite and to transmit passive immunity to individuals suffering from infectious disease (13,14). Purified non-catalytic antibody preparations with targeted specificity have been shown to favorably alter the pharmacodynamics of toxic compounds (15-17). The first catalytic monoclonal antibody was reported by Tremontano and his colleagues in 1986 (18). The technique depended on producing antibodies with specificity for a rationally designed structure representing the transition state for a modelled organic or enzymatic reaction. That first catalytic antibody hydrolyzed an aryl ester. Antibodies developed since then have catalyzed carbonate hydrolysis, amide hydrolysis, lactonization, aminolysis, and a Claisen rearrangement. The topic has been the focus of several recent reviews (19,20).

The data reported here, while preliminary and subject to confirmation as greater quantities of IIA12 ID10 become available, suggest that a catalytic antibody capable of hydrolyzing the highly toxic nerve agent soman can be developed. The k_{cat} of 4 min^{-1} for our catalytic antibody compares favorably with turnover numbers reported for other catalytic antibodies (19,20). Our results strongly support our selection of a trigonal bipyramid structure for a transition state analogue.

In an effort to minimize spontaneous base catalyzed hydrolysis there was no buffer included in this experiment. However, Broomfield, et al. (12) showed that in the absence of buffer hydrolysis was accelerated as the reaction proceeded due to the generation of hydrogen ions and a transition to acid catalyzed hydrolysis at the lower pH. We cannot be certain that the same phenomenon was not at work here. The outcome of this experiment under those circumstances would be worse than expected rather than better. Spontaneous hydrolysis would have increased, and the antibody catalyzed reaction would probably have slowed down. Therefore, we may see an increase in the k_{cat} when the reaction is repeated in a buffered system.

Due to a limited supply of the antibody it was not possible to include a second set of control samples containing heat inactivated antibody protein.

BRIMFIELD, LENZ, MAXWELL, BROOMFIELD

This would have ensured that the activity observed was not attributable to a non-specific function of the presence of protein in the experimental system. This is a valid concern. de Jong, et al. showed that serum proteins produce a finite increase in the hydrolysis of pure soman isomers in the absence of enzymatic activity (21). This effect will be addressed in future work.

Work with the catalytic antibodies has not yet reached the point where we can actually perform *in vivo* tests of their protective potential. However, we have been made very optimistic by the recent report of Broomfield's work (4) in which he successfully protected mice against the effects of twice the lethal dose of soman using a recombinant organophosphorus anhydride hydrolase from *Pseudomonas diminuta*. Our work is continuing with the use of additional protein conjugates that will allow us evaluate the effectiveness of other transition state analogues.

BIBLIOGRAPHY

1. C.A. Broomfield, D.M. Maxwell, R.P. Solana, C.A. Castro, A.V. Finger and D.E. Lenz. *J. Pharmacol Exp. Ther.* 259, 633 (1991).
2. B.P. Doctor, C. Castro, D. De La Hoz, A. Finger, G. Gold, H. Hively, D. Maxwell, S. McMaster, R. Solana, A.D. Wolf and C. Woodard. *FASEB J.* 5, A1599 (1991).
3. D.M. Maxwell and B.P. Doctor., in *Chemical Warfare Agents*, S.M. Somani, ed. (Academic Press, San Diego, CA, 1992), pp. 195-207.
4. C.A. Broomfield. *Pharmacol. Toxicol.* 70, 65 (1992).
5. A. Fersht, *Enzyme Structure and Mechanism* (W.H. Freeman & Co., New York, 1985) pp. 236-237.
6. R.M. Moriarty, J. Hiratake, and K. Liu. *J. Am. Chem. Soc.* 112, 8575 (1990).
7. J.W. Goding. *J. Immunol. Meth.* 39, 285 (1980).
8. M. Schulman, C.D. Wilde, and G. Kohler. *Nature.* 276, 269 (1978).
9. J.S. Little, C.A. Broomfield, M.K. Fox-Talbot, L.J. Boucher, B. MacIver, and D.E. Lenz. *Biochem. Pharmacol.* 38, 23 (1989).
10. D.P. Dumas, S.R. Caldwell, J.R. Wild, and F.M. Raushel. *J. Biol. Chem.* 264, 19659 (1989).

Photovoltaic Effect in Thin Ferroelectric Films for Nonvolatile Memory

Philip S. Brody, Kelly W. Bennett, and B. J. Rod

Harry Diamond Laboratories, U.S. Army LABCOM, Adelphi, MD 20783

1. Introduction

Recent developments in depositing ferroelectric materials, particularly lead zirconate-titanate (PZT), as thin films on silicon substrates have resulted in the prototype development of nonvolatile destructive-readout random-access memory (RAM). In such prototype devices, a silicon integrated ferroelectric (SIF) thin-film capacitor stores information in the form of a permanent or quasi-permanent remanent polarization.¹ These RAM's are intended to have on-chip densities and clock speeds comparable to that of conventional dynamic RAM's (DRAM's).

Nonvolatile memories are particularly attractive to military users and for space applications. Programmed information is more likely to survive an environmental upset due to radiation if the memory is nonvolatile. Additionally, radiation has a limited effect on remanent polarization within a PZT film,² and memory state loss due to radiation effects is expected to be minimal.

The fact that the films are thin allows the relatively low voltages available in the conventional complementary metal-oxide semiconductor (CMOS) and bipolar technologies to produce fields in the capacitor dielectric (ferroelectric) large enough to reverse the direction of the remanent polarization. In the destructive-readout memory,³ a voltage pulse applied to the capacitor switches the polarization, leaving a reverse-saturated polarization state. Because the polarization is saturated (or nearly so), an additional switching pulse of the same polarity results in a relatively small flow of displacement current, while switching with an opposite-polarity pulse results in a much larger flow, as the polarization is reversed. Detecting the magnitude of the integrated current on a series-integrating capacitor after a read pulse is the means of determining the original memory state of the capacitor.

The capacitor must be switched to write and switched to read. After a read, it must be switched again to restore the original polarization. A persistent problem in

the development of destructive-readout ferroelectric memory arises because the amount of polarization switched generally decreases (fatigues) with continual switching.⁴ The rapid accumulation of switching cycles that would occur in mostly read operations eventually leads to the loss of the capacitor's memory content. Current materials research efforts may reduce this fatigue problem. Whether these efforts will succeed is not certain at this time.

The loss of remanent polarization with increasing number of switching cycles would have less impact on memory reliability if the readout process were non-destructive, that is, if the polarization state could be determined without switching. A nondestructive readout has another advantage in that the programmed information is more securely stored, for it is maintained continuously in nonvolatile form. In a destructive-readout memory, the information is in a volatile form after reading and before the restore; during this period the stored information can be expected to be sensitive to environmental upset.

One approach to nondestructive readout has been the ferroelectric gate field-effect transistor (FET).⁵ Here the SiO_2 gate insulator of a FET is replaced by a ferroelectric film. The ferroelectric is then the dielectric in a capacitor in which one electrode is metal and the other a semiconductor. Switching the ferroelectric results in positive or negative screening charge in the semiconductor, where the effect is then an accumulation layer, or a depletion layer with the formation of an inversion channel. The resulting modulation of current in the channel identifies the memory state. These devices, however, are in only preliminary development. Long-time persistent channel modulation has not as yet been demonstrated.

Another possibility for SIF nondestructive readout depends on the use of optically generated currents in ferroelectrics, the direction and magnitude of which depend on the direction and magnitude of remanent polarization. The polarization-dependent photovoltaic phenomenon is well known in ferroelectric crystals and ferroelectric ceramics;⁶⁻¹⁰ it has also been observed in thin ferroelectric films.¹¹⁻¹³

We had demonstrated elsewhere^{14,15} the principle of nondestructive readout using photovoltaic currents. In this earlier work, a ferroelectric capacitor was formed as an electroded region on a ceramic ferroelectric substrate. Photovoltaic current flowing from the capacitor and integrated as charge on the gate of a FET produced a gate voltage that modulated the drain-source channel current, the modulation depending on the direction of the capacitor polarization. Moderately high-speed write, nondestructive read, and long-term memory retention were shown. This development work at Harry Diamond Laboratories was partially supported by NASA-Langley; the major interest was for space applications.

It would be difficult, however, to make practical a memory in which the ferroelectric capacitor is on a bulk ceramic substrate. Practicality requires a silicon integrated form using a thin-film ferroelectric capacitor with adequate polarization-dependent photovoltaic response; additionally, electrodes must be sufficiently transparent to transmit the required optical flux. In this respect, there is a lack of sufficient experimental data available; also the photovoltaic behavior of the thin films differs considerably from that of the bulk ceramics used in most previous studies of the photovoltaic phenomenon in ferroelectrics. The behavior in films is not presently well understood.

We have proceeded therefore to further investigate photovoltaic phenomena in ferroelectric films. In this paper, we describe an experimental investigation of photocurrents from two different lead-based perovskite oxide films, particularly under conditions of polarization reversal. We discuss a possible mechanism that explain the photovoltaic currents, describing a model which yields an internal potential distribution following polarization reversal and from which the source of the currents and their time dependence can be understood. We discuss potential applications of the polarization-dependent photovoltages to memory and processing.

2. Photovoltaic Phenomena in Ferroelectrics

Anomalous steady-state photovoltages, dependent on remanent polarization, are well known in bulk ferroelectric ceramics; they are also found in ferroelectric single crystals.⁶⁻¹⁰ The photovoltages are considered anomalous because they cannot immediately be ascribed to an ordinary photovoltaic mechanism, such as band bending at metal contacts or the Dember effect; both these mechanisms are photo-effects where the open-circuit photovoltages (photo-emfs—electromotive forces) are limited to the potential difference across the band-gap (about 3 V in the lead-based perovskites). Photo-emfs in ferroelectrics, however, often exceed the bandgap potential difference, with steady open-circuit voltages observed exceeding even 1000 V across illuminated bulk ceramics and single crystals. The phenomenon producing the emfs in single crystals is thought to be a true bulk effect, which is produced by free carriers from localized states in the presence of microscopic asymmetry at the absorbing centers.^{8,9} In the ceramics, however, the high-voltage emfs are thought more likely to result from a series summation of emfs originating at the boundaries of the ceramic grains. The grain boundary mechanism is thought to be a surface or contact effect, as in the ordinary photovoltaic mechanism.^{9,16}

Photovoltaic currents are also seen in ferroelectric thin films.¹¹⁻¹³ The effect in the films appears to be a surface or contact phenomenon, as in the ceramics, rather than a true bulk effect, as in some monolithic crystals.

3. Experimental Studies of Photovoltaic Currents

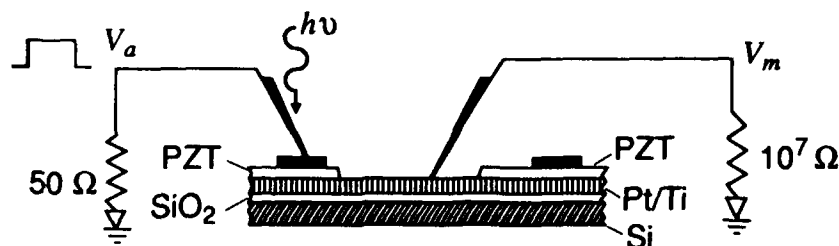
In our experimental study, we used two different sol-gel process lead-based perovskite films:¹⁷ a lead zirconate-titanate (PZT) film with a 53/47 zirconium-titanium composition ratio and a lead titanate (PT) film. The films were on platinum-coated silicon wafers. Thin-film capacitors, $150 \times 150 \mu\text{m}$ square, were formed by sputter-depositing top platinum electrodes using a micromachined silicon shadow mask.¹⁸ The electrodes were thin (about 200 \AA thick) to allow transmission of illumination. Contact to top electrodes and the base electrode was made with probes (fig. 1), using a probe station. Light from a mercury arc lamp passing through the central port of the trinocular probe station microscope illuminated the capacitors. The electrodes transmitted typically about 15 percent of incident illumination.

The illumination could be filtered with a glass absorption filter placed in the optical train; the filter transmitted wavelengths in a band between 320 and 390 nm. By taking into account the filter attenuation, one sees that photovoltaic currents result almost exclusively from this near-ultraviolet band of wavelengths. This is heavily absorbed light from absorption edge wavelengths (about 370 nm in these lead perovskite oxides). The illumination intensity of filtered light at the ferroelectric film surface after it passed through the electrode was typically about 2.4 mW/cm^2 .

The top electrode (fig. 1) was connected to ground through a $50\text{-}\Omega$ resistor. Voltage pulses at V_a switched the remanent polarization of the capacitor. The lower electrode was connected to ground through a chosen resistance between 10^6 and $10^{11} \Omega$. It was also connected to the input of an electrometer amplifier; the output of the electrometer went to a chart recorder.

The photovoltaic current flowing through the output resistance is the voltage at the electrode divided by the resistance to ground. Photovoltaic currents were independent of the load resistor and proportional to the illumination intensity for lower resistance values (short-circuit currents). There was thus a characteristic open-circuit voltage. Current outputs were a function of load resistance, as would be explained by an equivalent circuit like that seen in ceramics^{6,7} in which the emf appears in series with a photoresistance. This was only approximately so; there were clear indications that the photoresistance depended weakly on the direction of the photo-

Figure 1. Arrangement for contacting capacitor electrodes to switch remanent polarization and measure photovoltaic currents.



emf. The values of the steady open-circuit voltage and short-circuit current differed for the two films; these are shown in table 1.

The results shown in table 1 are steady values for saturation negative polarization and unfiltered mercury arc illumination. Introduction of a Corning CS 7-39 2-mm-thick filter, with a filter factor of about 0.5, reduces the illumination intensity to 2.4 mW/cm^2 of absorption edge illumination. The short-circuit current is also reduced by about 50 percent; the open-circuit voltages are just marginally reduced. It thus appears that the photocarriers are generated largely from absorption edge illumination.

The photovoltaic currents from the PT film were more than an order of magnitude greater than those from the PZT films; the open-circuit voltage was somewhat less. A factor contributing to this result (which is similar to what is observed in ceramic materials) may be the higher remanent polarization of the lead titanate.¹⁰ The steady values in table 1 were characteristic not only of the saturation negative remanent polarization, but also of the capacitors in the as-fabricated state before a switching voltage was applied. This implies that the as-fabricated capacitor was in a saturation negative remanent polarization state.

Results for photovoltaic current versus time during and immediately after the polarization was switched with the capacitor illuminated are shown in figure 2 (PZT film) and in figure 3 (PT film). The switching pulses were $\pm 8 \text{ V}$ and $\pm 10 \text{ ms}$ long.

There is characteristically an initial state photovoltaic current. When a positive switching pulse is applied, the photovoltaic current increases markedly; this peak current then decays rapidly to a new steady value. Switching the polarization again, with a pulse of the opposite polarity, produces a second, oppositely directed peak current. This negative peak current again decays quickly to a second steady-state value.

The steady values of photovoltaic currents after switching are shifted with respect to the values before switching. The currents are not shifted in a symmetric

Table 1. Open-circuit voltage and short-circuit current for unfiltered mercury arc illumination of 44 mW/cm^2 after transmission through electrode

Note: values are for saturating negative remanent polarization.

Sample	Open-circuit voltage (V)	Short-circuit current (nA)
PZT (53/47)	-0.52	-0.45
PT	-0.30	-7.5

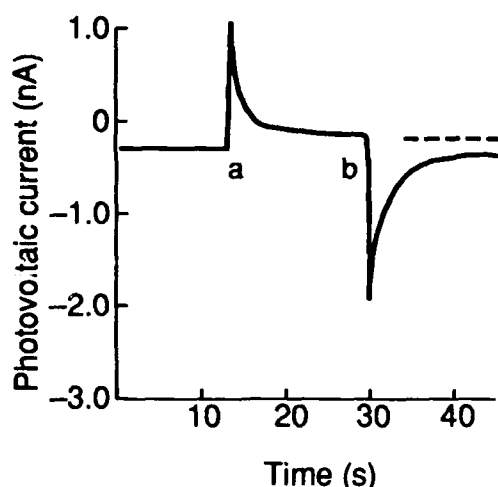


Figure 2. Photovoltaic current as a function of time for illuminated capacitor upon polarization reversals. Film is 0.26- μm -thick 53/47 Zr/Ti ratio PZT, with (a) positive pulse applied and (b) negative pulse applied.

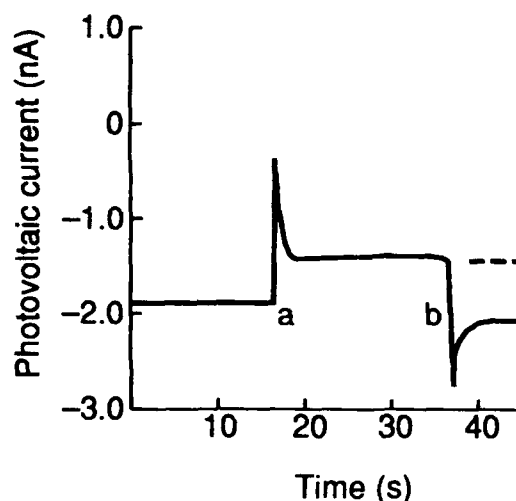


Figure 3. Photovoltaic current as a function of time for illuminated capacitor upon polarization reversals. Film is 0.35- μm -thick PT, with (a) positive pulse applied and (b) negative pulse applied.

way with respect to zero current; the asymmetry is the result of a constant photovoltaic current component that is independent of the direction and magnitude of the switched polarization. The polarization-dependent shifts are linearly proportional to the pulse-induced change in remanent polarization.¹²

The magnitude of the change in the steady photovoltaic current after polarization reversal decreases slowly with time, decreasing linearly with the logarithm of elapsed time. Figure 4 shows the slow logarithmic portion of the decay for the PT film, where the remanent polarization had been switched to saturation positive (curve 1) and saturation negative (curve 2) levels with multiple 100-ms 8-V pulses. The dashed line represents the current before any pulse is applied and is the polarization-independent current. We note that the currents for positive and negative polarization are roughly symmetric with respect to this level.

4. Mechanism for Photovoltaic Currents

There are several possible sources for the component of the photovoltaic current that is independent of both the direction and magnitude of the remanent polarization. These include the contact potential at the metal/ferroelectric interface and other interface potentials—for example, potentials resulting from charge accumulated in surface states.¹⁹ These potentials would be expected to produce fields of opposite polarity on opposing surfaces. They are a possible source of thin surface

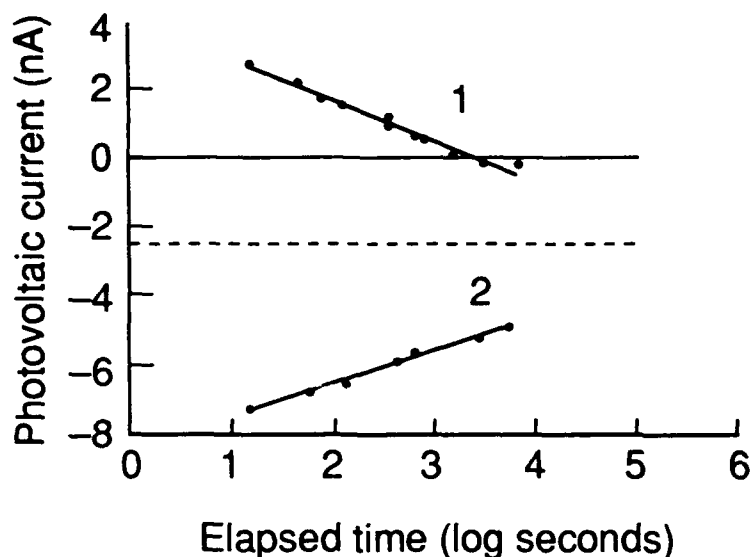


Figure 4. Slow decay of steady photovoltaic current from a lead titanate film capacitor for positive (curve 1) and negative (curve 2) saturation remanent polarization. Curves intersect at about 250 hours.

layers on ferroelectric crystals and ceramic grains (Kanzig exhaustion layers²⁰), where high fields from these potentials might be expected to polarize the material and saturate the dielectric susceptibility. These fields might also be expected to prevent polarization reversal in the surface regions. Photovoltaic currents produced as a result of such surface space-charge fields might be expected at first observation to cancel each other. The photocarrier-generating illumination is, however, strongly absorbed and creates more carriers at the top interface than at the substrate electrode interface. The potentials then act on unequal charge densities, and there is a flow of current.

We now consider a mechanism for the rapidly decaying transient photovoltaic current initially present after switching and the final polarization-dependent steady current that follows. We believe that both current components (the rapidly decaying initial current and the slowly decaying steady current) result from a time-varying potential distribution within the film, which is produced as remanent polarization is screened by mobile charges generated by the illumination. These charges move to screen an initial depolarization field.

If all the compensating charge hypothetically is distributed in a plane exactly coincident with the plane of the polarization discontinuity, there is no field within a film. If the termination of the remanent polarization is separated from the compensating charges (which might occur if the compensating charge were on the electrodes and there were a nonswitching surface layer), there is then a non-zero potential distribution within the film.^{16,21} The fields from such a potential distribution can be expected to generate photovoltaic currents; the polarity of these photovoltaic

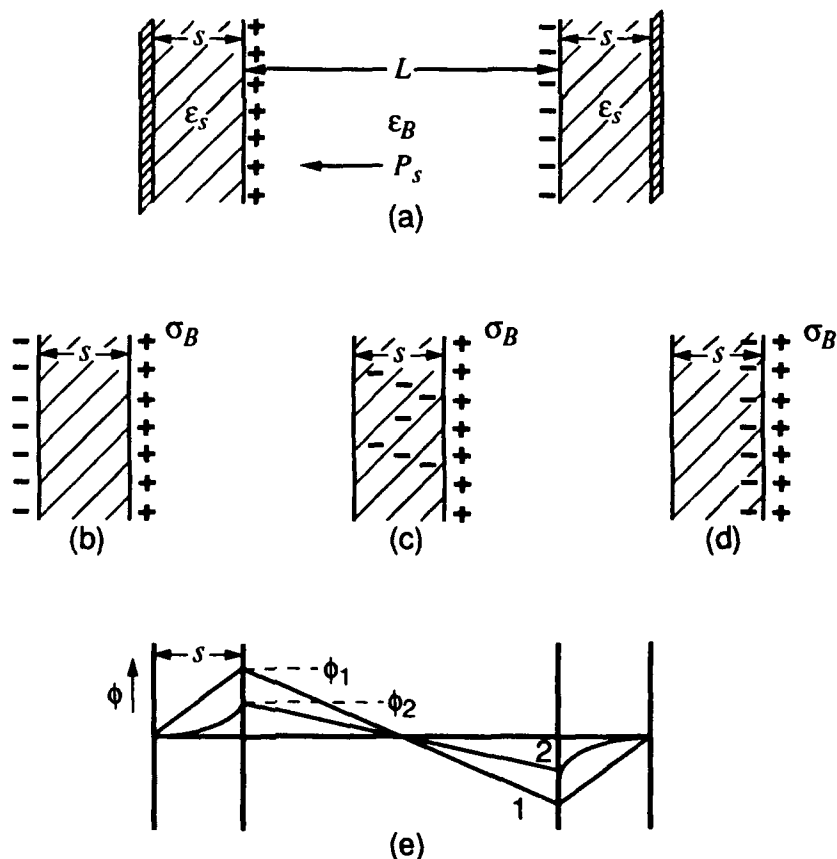
currents will depend on the direction and magnitude of the remanent polarization. We consider a simple model in which the polarization switches within the bulk of a film, but not in thin surface layers.

5. Potential Distribution in Films due to Nonswitching Surface Layers

The initial potential distribution in the film after polarization reversal is calculated in terms of the parameters of a simple model (fig. 5a) of a thin film with conducting electrodes.

The polarization may be reversed only within the bulk of the film in the region of thickness L , but not in the surface regions of thickness s (the surface regions may be polarized, but the polarization is not reversible). After polarization reversal there is a large depolarization field $-P_s/\epsilon_0\epsilon_B$ from the reversed bound charge $\sigma_B = P_s$, where ϵ_0 is the permittivity of free space. The depolarizing field is compensated initially by charge on the electrodes (by conduction through the pulse generator output impedance). The resulting potential distribution (fig. 5e, curve 1) is calculated from Gauss's law; we have assumed that the median potential is always zero

Figure 5. Model for thin film: (a) Structure of film of thickness $2s + L$; remanent polarization is P_s , surface layer thickness s , surface layer relative dielectric constant ϵ_s , and film bulk relative dielectric constant ϵ_B . Examples of possible distributions of compensating charges which screen polarization (b) on electrodes, (c) uniformly distributed in surface layers, and (d) near discontinuity in polarization. (e) Potential distribution ϕ_1 from charge distribution b and ϕ_2 from distribution c (uniform within surface layer).



from symmetry considerations. The potential has a maximum and minimum at the surface layer-bulk interfaces whose magnitudes are

$$\phi_1 = P_s \frac{s}{\epsilon_0 \epsilon_B \left(1 + \frac{\epsilon_B s}{\epsilon_s L} \right)} \quad (1)$$

Brody and Crowne¹⁶ previously calculated the potential difference across a surface layer of thickness s for the case of shorted electrodes and compensating charge uniformly distributed within the layer (fig. 5c). In this case, the potential distribution across the surface layer has a different form (fig. 5e, curve 2), with a maximum and minimum of half those calculated for the case in which the entire screening charge is confined to the electrodes (curve 1). If the screening charge is not uniformly distributed within the layer, but confined only to a narrower region very close to the polarization discontinuity, then the potential drop across the bulk is further reduced, approaching the zero potential everywhere. That is the case only for the screening charge precisely juxtaposed on the termination of the polarization vectors.

When the screening results from equilibrium carriers, the maximum potential drop across the surface layer is limited to the bandgap potential difference.²² Because of the low conductivity of the perovskite oxides, the time constant for such screening may be long.

The effect of a nonequilibrium carrier density produced by illumination (photocarriers) on the potential distribution can now be used to explain the photovoltaic currents from the film capacitor. The proposed mechanism is the same as that used to explain photo-emfs across ceramic grains.¹⁶ The photocarrier density created by the strongly absorbed bandgap illumination is assumed sufficient to screen the relatively weak field within the bulk but not the strong fields in the surface regions. Figure 6 shows the open-circuit potential distribution which results from such screening when the compensating charge is confined to the electrodes.

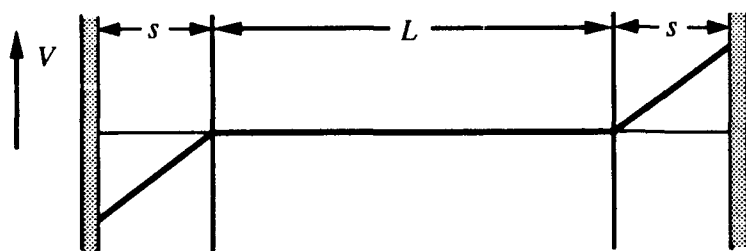


Figure 6. Open-circuit potential distribution in illuminated film. Voltage difference across film is $\delta V_{oc} = 2\phi_1$. This is for case where charges compensating polarization charge are restricted to electrodes.

In this proposed explanation for the photovoltages, an initial potential distribution in the dark after a polarization reversal (shown in fig. 5e) produces no voltage across the electrodes, even when the circuit is opened. The screening length in the dark is very large (because of the very small number of equilibrium carriers), and a potential drop exists across the bulk, which cancels the potential drop across the surface layers. Illumination produces photocarriers which screen the relatively weak internal fields due to the potential drop across the bulk, leaving only the high surface fields due to the potential drops across the surface layers. If we assume that the charge distribution is unmodified by the capacitor being shorted, then the short-circuit current is the drift current in the open-circuit potential. For a simple photoresistance in the bulk, the voltage V_m across the resistive load R_L is given in terms of the open-circuit voltage and the photoresistance R_{ph} by

$$V_m = \frac{V_{oc}}{\frac{R_{ph}}{R_L} + 1} \quad (2)$$

A screening length (Debye length) for nonequilibrium carriers¹⁹ is given by

$$L_D = \sqrt{\frac{\epsilon_0 \epsilon_r kT}{e^2(n_0 + p_0)}} \quad (3)$$

If the screening length is less than the characteristic length of the potential variations, those potential variations and the resultant fields are essentially eliminated. The screening length for the bulk thus must be significantly less than the thickness of the bulk portion of the film.

For the screening length to be less than 0.1 μm requires a carrier density, $n_0 + p_0$, of $1.1 \times 10^{17} \text{ cm}^{-3}$ for the PZT film and of $1.7 \times 10^{16} \text{ cm}^{-3}$ for the PT film using the bulk values for the dielectric constants, $\epsilon_r = 800$ for the PZT film and $\epsilon_r = 120$ for the PT film. The values of carrier density could be obtained from carrier lifetime, quantum efficiency and absorption depth;²³ there is presently, however, only limited information on this topic and only for the perovskite oxide BaTiO_3 .^{24,25} However, the above values are reasonable numbers for photogenerated carrier densities.

The time dependence of the polarization-dependent photovoltaic currents originates, we believe, in the following way. The initial peak photovoltaic current results from illumination-generated charge screening the potential across the bulk of

the film. The potential distribution in the dark is that resulting from an initially charge-free surface layer. The switched polarization is screened by charge entirely confined to the conducting electrodes (see fig. 5b and the potential distribution curve 1, fig. 5e). This current decays because the potential difference across the gap producing the current is reduced as mobile charges created by the illumination are attracted to and populate trap sites within the surface layer (fig. 5c, 5d). The result is a final almost steady equilibrium potential distribution, where the charge compensation or screening is from trapped charges distributed within the surface regions (fig. 5c, 5d, and potential distribution curve 2, fig. 5e). The fields across the surface layers drive the photovoltaic current, and as these fields decrease, the current decays. This accounts for the rapid decay in the photovoltaic current. The potential distribution now approaches a steady state. Fields from the steady-state distribution drive a steady photovoltaic current. The slow decrease in this current, which follows linearly the logarithm of elapsed time (fig. 4), appears to be a result only of a slow decay in the remanent polarization. The surface layer fields and the photovoltaic currents that they drive decrease in time, because the remanent polarization is slowly decaying.

The possible mechanism for the decay of remanent polarization includes the action of transverse stress, or the action of the residual field in the bulk region causing a slow depolarization. The bulk depolarization field, we note, is measured simply as the open-circuit photovoltage divided by the film thickness.

6. Calculation Estimating Surface Layer Thickness and Internal Fields

We have assumed in the discussion that the surface layers are very thin in comparison to the thickness of the film. We can use equation (1) to obtain an estimate of the surface layer thicknesses that are consistent with our assumption of photo-induced nonequilibrium carrier screening and the measurements of photovoltaic current. The value ϕ is in principle the photo-emf that appears immediately after an illuminated capacitor is switched. It cannot, however, be obtained directly from a measurement of open-circuit voltage, because the time required to charge the ferroelectric capacitor under open-circuit conditions exceeds the decay time of the initial photovoltage. Instead, we estimate this value from the easily measured steady photovoltage and the ratio between the peak photovoltaic current and the steady photovoltaic current. Because the emf and currents are proportional, the initial photo-emf is the ratio between the initial and steady photovoltaic currents multiplied by the steady open-circuit photovoltage. We also use the value of the film thickness and estimates of the dielectric constants in bulk and surface. Using bulk ceramic values for the dielectric constant in both the film bulk and in the film surface

layers, we find that in the PT film the surface layer is 11 Å thick and in the PZT film it is 75 Å thick. The surface layer thickness depends on the estimate of the surface layer dielectric constant, which can be expected to be reduced by the saturating effects of the high fields in the surface region. Assumptions of lower surface layer dielectric constants will reduce the calculated thicknesses. The surface region fields, $E_s = \phi_{max}/s$, are thus very large: they considerably exceed coercive fields for these materials and may be expected to produce oppositely directed remanent polarizations in the surface regions, which cannot be switched.

7. Potential Application to Nonvolatile Nondestructive-Readout RAM and to Multiplicative Processing

Because the photovoltages from the ferroelectric film capacitors are strongly dependent on remanent polarization, with the photovoltages shifting when the polarization is reversed, they may be applied to nonvolatile nondestructive-readout RAM. We note that although the shift in photovoltage initially decreases rapidly, the direction of the shift depends only on the direction of the remanent polarization change. Thus the photovoltage could be applied to the gate of a FET to modulate the channel current,^{14,15} and we could depend on the direction of the shift in the photovoltage with polarization reversal to cause an increase or a decrease in the channel resistance. This change would then be used to ascertain the memory state, as in the ferroelectric FET concept.⁵

Such a steady portion of the photovoltages on the gate can be expected to remain constant, decaying only as the result of a slow loss of switched remanent polarization due to aging, as in the case of the PT film (fig. 4). The retention loss in the PT film is clearly abnormally high; retention loss for other compositions can be expected to be less.

Producing the voltage on the gate requires that the ferroelectric capacitor charge to its open-circuit voltage as the result of illumination. For a given illumination intensity, the capacitor will charge in a time that is independent of its size. The internal photoresistance is inversely proportional to the capacitor area, and the capacitance is directly proportional to the capacitor area; the result is an RC charging time independent of area. This charging time is simply the product of photoresistivity and dielectric constant. For the lead titanate film illuminated at 2.4 mW/cm², the charging time is about 3 ms. The charging time for the PZT is much larger, because of the higher photoresistivity and larger dielectric constant. The calculated value is 160 ms. We note that a thin-film capacitor of very small size will charge the very low gate capacity of a FET to the open-circuit voltage at a rate given by the calculated RC time. Only for very large gate capacitance will the capacity of the gate load the charging circuit and decrease the charging time.

Once charged, the voltage on a gate need only be maintained by sufficient current to compensate for leakage, and once the gate is charged, channel current can be interrogated at any time to reveal the polarization state of the ferroelectric capacitor.

It is not necessary to wait for a capacitor to charge before reading during high-speed read/write applications, if cell circuitry is used in which the write pulse simultaneously switches the polarization and charges the gate.¹⁵

We now note that the switched short-circuit photovoltaic current depends bilinearly on the switched remanent polarization and the illumination intensity. This makes possible another application, in addition to nonvolatile nondestructive-readout RAM. This application is an analog multiplicative processor, in which one factor in a product is partially encoded remanent polarization and the other is illumination intensity. Such multiplication can be used, for example, in image correlation²⁶ for image recognition.

Acknowledgements

We thank Dr. S. Bernacki of Raytheon Corporation and Dr. S. W. Dey of Arizona State University for providing the films used in this study.

References

1. J. F. Scott and C. A. Paz de Aroujo, *Science* **240** (1989), 1400.
2. J. M. Benedetto, R. A. Moore, F. B. McLean, P. S. Brody, and S. K. Dey, *IEEE Trans. Nucl. Sci.* **27** (1990).
3. J. T. Evans and R. Womack, *IEEE J. Solid State Circuits* **23** (1988), 1171.
4. See for example, H. M. Duiker et al, *J. Appl. Phys.* **68** (1990), 5783.
5. S. Y. Wu, *IEEE Trans. Electron Devices*, **ED21** (1974), 499.
6. P. S. Brody, *Solid State Commun.* **12** (1973), 673.
7. P. S. Brody, *Solid State Chem.* **12** (1975), 193.
8. A. M. Glass, D. von der Linde, and T. J. Negran, *Appl. Phys. Lett.* **25** (1974), 234.
9. V. M. Fridkin, *Photoferroelectrics*, Springer Verlag, New York (1979), p 103.
10. K. Uchino, Y. Miyawaza, and S. Nomura, *Jpn. J. Appl. Phys.* **2** (1982), 1671.
11. V. S. Dharmadhikaru and W. W. Grannemann, *J. Appl. Phys.* **53** (1982), 1671.
12. P. S. Brody and B. J. Rod, presentation at the 3rd International Symposium on Integrated Ferroelectrics, Colorado Springs, CO, 1991, preprint HDL-PP-NW-91-4; available from the authors (to be published in Proceedings).
13. S. Thakoor, A. Thakoor, and S. Bernacki, presentation at 3rd International Symposium on Integrated Ferroelectrics, Colorado Springs, CO, 1991 (to be published in Proceedings).

BRODY, BENNETT, ROD

14. P. S. Brody, *Appl. Phys. Lett.* **38** (1981), 673.
15. P. S. Brody and E. Katzen, *Nonvolatile Memory with RAM Capabilities and Intrinsic Radiation Hardness*, Harry Diamond Laboratories, HDL-TM-82-14 (1982).
16. P. S. Brody and F. Crowne, *J. Electron. Mater.* **4** (1975).
17. K. D. Budd, S. K. Dey, and D. A. Payne, *Br. Ceram. Soc. Proc.* **36** (1985), 107.
18. B. J. Rod, *Preparation of Ferroelectric Samples for Electrical and Radiation Characterization Studies*, Harry Diamond Laboratories, HDL-TM-91-14 (1991).
19. M. Many, Y. Goldstein, and N. B. Grover, *Semiconductor Surfaces*, North Holland, Amsterdam (1965), pp 5, 157.
20. W. Kanzig, *Phys. Rev.* **98** (1955), 549.
21. P. Wurfel and I. F. Batra, *Ferroelectrics* **12** (1976), 55.
22. G. M. Guro, I. E. Ivanchik, and N. F. Koutonyuk, *JEPT Lett.* **5** (1967), 5.
23. S. M. Ryvkin, *Photoelectric Effects in Semiconductors*, Consultants Bureau, New York (1969), pp 1-18.
24. M. Cardona, *Phys. Rev.* **140** (1965), A651
25. H. Tredgold, *Space Charge Conduction in Solids*, Elsevier (1966), p 79.
26. P. S. Brody, *Photovoltaic-Ferroelectric Correlation Devices*, U.S. Patent 4,126,901 (1978).

Scene Segmentation and Definition for Autonomous Robotic Navigation Using Structured Light Processing

Dr. C. David Brown
U. S. Army Combat Systems Test Activity
Aberdeen Proving Ground, MD 21005-5059

1. Introduction

Substantial research has been directed toward the application of machine vision to the problem of the guidance of robots and autonomous vehicles along pathways that are not previously designated by special markings of some kind [1-3]. The primary function of the machine vision system is to provide the autonomous robotic navigation system with information describing the location and description of the pathway in front of the vehicle. Color, and its closely associated gray level contrasts, is a commonly employed method of obtaining pathway location, however, during off road navigation, many instances can be imagined where the color of the pathway differs little from that of the background. In these situations, the three dimensional (3-D) texture, or change in shape, in the scene must be sensed by the visual system, and a region of low texture is designated as pathway, and high texture as background.

Stereo vision employing two video cameras, closely models the human vision system for sensing the 3-D characteristics of a viewed scene. However, the processing is complicated by the problem of determining pairs of corresponding points in the two images. Researchers have investigated replacing one of the cameras by a structured light source, such as projected lines or a grid, for autonomous vehicular navigation [1], and this configuration is becoming commonplace in industrial machine vision applications [4]. First, the correspondence problem is eliminated because each element of the structured light is projected at a known angle and sensed as a definable spot or shape in the camera image. Second, only the image of a single video camera needs to be processed.

Additional advantages are realized if a laser is used as the projection source. A laser is a highly collimated, high power source. Therefore, the portions of the scene illuminated by the laser contrast well with portions not illuminated. Because the laser is also monochromatic, narrow band optical filters can be employed to enhance this contrast. Therefore, laser projected structured light yields an image which is easily thresholded for further processing. The thresholded image then becomes a binary, two dimensional (2-D) representation of the 3-D texture in the scene, which is then conducive to simplified regional processing to segment the image based on the 3-D texture.

The real advantage of this processing method is that critical processing parameters are defined in terms of 3-D scene measurements, but processing is accomplished almost entirely in 2-D, thus greatly simplifying the processing. For example, acceptable vertical pathway texture should be something less than vehicle ground clearance. Given predefined sensing system parameters, such as camera location and lens focal length and laser location and scan angles, these scene measurements can be projected into the 2-D processing plane and stored in a look-up table to facilitate processing.

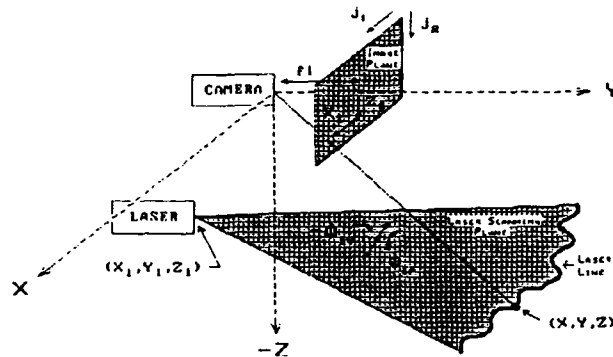


Figure 1: 3-Dimensional geometry

When autonomous vehicles are navigating along predefined roadways, it is a reasonable assumption that the grade and cant of the roadway will not exceed the vehicle's capabilities. However, with off-road navigation or planetary exploration, no such assumption can be made concerning pathways chosen by the autonomous vision system through an unstructured environment. Therefore, the vision system must also have the ability to define the 3-D characteristics of the chosen pathway to determine its acceptability. Further processing of the image of a structured light grid, projected on the scene ahead of the autonomous vehicle, can accomplish the required pathway definition. Thus, this method may provide a valuable augmentation to present autonomous vehicular vision systems to accomplish the difficult task of off-road navigation or planetary exploration.

2. Background

Examination of Fig. 1 demonstrates the geometric principles upon which the scene segmentation method is based. The camera image of a horizontal scan line changes in response to the 3-D texture of the scene. The 3-D texture of the pathway differs greatly from that of the background in the scene, with the pathway texture being considerably more fine and regular. Therefore, the shape of the grid lines varies significantly between the pathway regions and background regions of the image. The application of the above descriptions of image texture and the accompanying segmentation methods provides a first order processing method to identify the pathway boundaries for input to a route planner for an autonomous robot or vehicle.

If adequate geometric parameters are known about the structured light projection and the camera sensing system, then the 3-D characteristics of the scene can be determined [5]. The camera focal point can be chosen as the origin, with the laser scanner focal point at (x_l, y_l, z_l) . The laser scan rotation angle (ϕ_{lr}) and vertical projection angle (ϕ_{lv}) are known, and the camera image location (x_i, f_l, z_i) is calculated based on the previously described camera calibration. The coordinates (x, y, z) of the laser spot in the scene can then be derived by calculating the point of intersection of the camera ray:

$$\frac{x}{x_i} = \frac{y}{f_l} = \frac{z}{z_i} \quad (1)$$

with the laser scanning plane:

$$\frac{x}{x_0} + \frac{y}{y_0} + \frac{z}{z_0} = 1 \quad (2)$$

where the planar intercepts are:

$$x_0 = x_l + y_l \tan(\phi_{lv}) \sin(\phi_{lr}) + \frac{z_l}{\tan(\phi_{lr})} \quad (3)$$

$$y_0 = \frac{\sqrt{x_l^2 + z_l^2} \sin(\phi_{lr} + \tan^{-1}(\frac{z_l}{x_l}))}{\tan(\phi_{lv})} + y_l \quad (4)$$

$$z_0 = x_l \tan(\phi_{lr}) + y_l \tan(\phi_{lv}) \cos(\phi_{lr}) + z_l \quad (5)$$

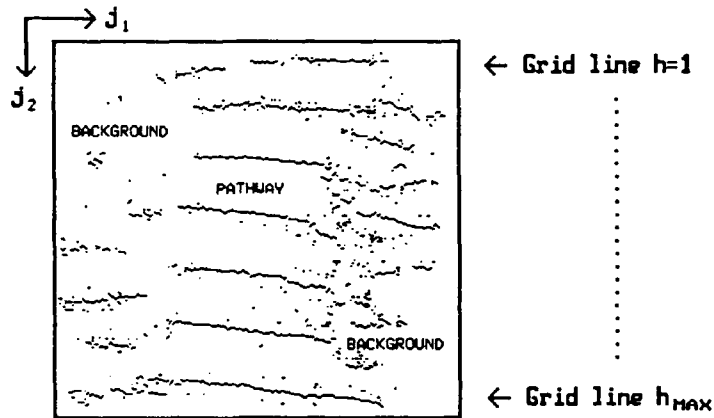
Solving for x, y , and z yields:

$$\begin{bmatrix} \frac{1}{x} \\ \frac{1}{y} \\ \frac{1}{z} \end{bmatrix} = \begin{bmatrix} 1 & \frac{f_l}{x_l} & \frac{z_l}{x_l} \\ \frac{x_l}{f_l} & 1 & \frac{z_l}{f_l} \\ \frac{x_l}{z_l} & \frac{f_l}{z_l} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{x_0} \\ \frac{1}{y_0} \\ \frac{1}{z_0} \end{bmatrix} \quad (6)$$

The laser scanner traces an approximately parallel grid on a planar surface normal to the projection axis. This is an essential characteristic of the horizontal scan lines, so that the segmentation procedure can be based on the detection of straight lines in the camera image. Also, the laser horizontal projection angle is not needed to derive the 3-D coordinates of the point in the scene, therefore vertical scan lines are not required in the structured light grid.

Each scan line is formed as the laser beam sweeps through a plane in space. The vertical projection angles associated with each horizontal line on the structured light grid are those corresponding to the horizontal planes swept by the laser beam as it traces the grid lines. As long as the grid scanning remains unchanged, a look-up table method can be utilized to obtain the laser vertical projection angle and rotation angle corresponding to a particular grid line of interest. In addition, the camera image location as shown in Fig. 1, is calculated based on a previously derived camera calibration and the location of the pixel in the camera image.

Utilization of extracted 3-D information from a scene to determine surface contour can be accomplished [6], and this becomes a valuable consideration in off-road autonomous navigation. Because a path for travel is desired, a significant concern is the planar content of a possible path. As explained previously, it is possible to stereoscopically locate various points throughout a surface of interest. Restricting interest to surfaces of a single planar orientation suggests that a collection of points residing on a surface can indicate the planar nature of that surface. It is likely that stereoscopically extracted points are not coplanar; thus, a description of the surface requires that a plane of best fit be determined. This can be accomplished by implementing a mathematical minimization technique (see Appendix), with the assumption that the surface under analysis consists of a single plane.

Figure 2: Image field, $t(j_1, j_2)$

3. Image Analysis

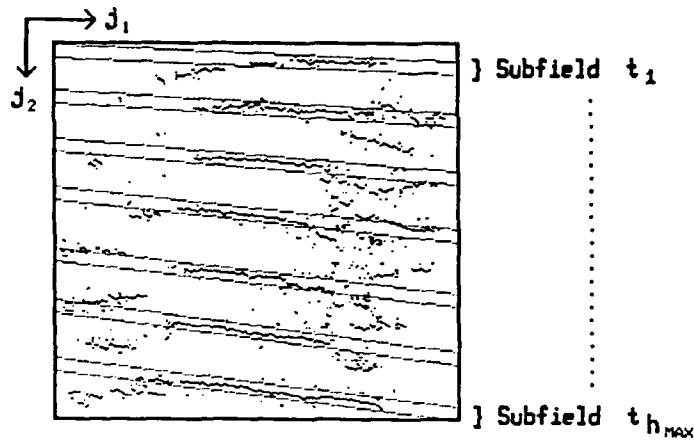
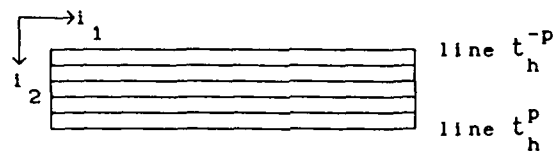
The captured video image of structured light horizontal grid lines represents a two-dimensional grey level image field denoted as: $a(j_1, j_2)$, where j_1 and j_2 are the horizontal and vertical spatial indices of each pixel in the field, and $a(j_1, j_2)$ is the grey level at that spatial image location. Because all desired information in the image is given by the fact that pixels are either illuminated by laser light or not, a binary image is desired to simplify further processing. As explained in the introduction, the image of the structured light is easily thresholded because of the high power density of laser illumination. The resulting binary image field is denoted as $t(j_1, j_2)$, (see Fig. 2).

A. Extraction of Processing Subregions

The 3-D texture in the scene is represented by the 2-D shape of the images of the laser lines. Therefore, considerable reduction in processing can be accomplished by only processing subregions of the image containing the laser lines.

The 3-D texture of regions in the scene is proportional to the spatial deviation of the image of the structured light grid lines. Pathway determination amounts to detection of lines in the thresholded image field. Line detection is accomplished by means of the Hough transform by mapping each image point into a parameter space based on the linear equation $j_2 = aj_1 + b$ where j_2 is the vertical index, j_1 is the horizontal index, a is the slope of the line, and b is the vertical intercept [7]. Alternately, the transformation can be based on the normal form of the linear equation, $j_1 \cos \theta + j_2 \sin \theta = \rho$, where ρ is the normal distance from the origin to the line, and θ is the angle from the horizontal axis to the normal to the line [7]. Acceptable pathways differ little from level, therefore, the grid line images are near horizontal, and the range of slopes of the lines is small and close to zero. Therefore, the Hough transform based on the linear equation $j_2 = aj_1 + b$ is appropriate and more computationally efficient.

The Hough transform maps every white (binary pixel value of 1) image pixel into a line in the (a, b) parameter space output of the Hough transform using the equation $b_h = j_2 - a_h j_1$. The range of slopes corresponds to acceptable pathway cant, thus excluding pathways with excessive cant. The range of slopes is quantized and sequenced, forming a line in the (a, b) plane corresponding to the (j_1, j_2) image spatial coordinates of each white pixel in the image field. The Hough transform output

Figure 3: Subregions denoted on image field, $t(j_1, j_2)$ Figure 4: Image subfield, $t_h(i_1, i_2)$

of a complete image amounts to the addition of the lines associated with each white image pixel. Peaks in the Hough transform parameter space, formed at points where many of these lines intersect, denote strong line tendencies in the image field, $t(j_1, j_2)$. Parameter space coordinates (a_h, b_h) are associated with each peak denoting a line with slope a_h and intercept b_h in the image field [11].

Image subfields are then extracted from the image field (see Fig. 3) where the center line (t_h^0) of each subfield, $t_h(i_1, i_2)$, corresponds to the mapping: $i_1 = j_1$ and $i_2 = a_h j_1 + b_h$, (see Fig. 4). Image lines of slope a_h are now horizontal in the subfield, $t_h(i_1, i_2)$. Each subfield is $2p + 1$ image lines high, such that eachline (t_h^l) in a subfield (t_h) corresponds to the mapping:

$$i_1 = j_1 \text{ and } i_2 = a_h j_1 + b_h + l, \quad -p \leq l \leq p. \quad (7)$$

B. Morphologically Based Image Segmentation

The extraction of information from the structured light image involves detection of the image location of a definable shape, such as a contiguous line for sensing texture. Morphology involves the design of an operator to sense this definable shape [8], and therefore is well suited to this application. Morphology is also very computationally intensive. However, because of its wide general applications, hardware implementations of morphological operations are being designed. The increased speed of these implementations makes the use of morphology very attractive for this application.

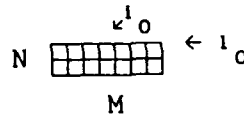
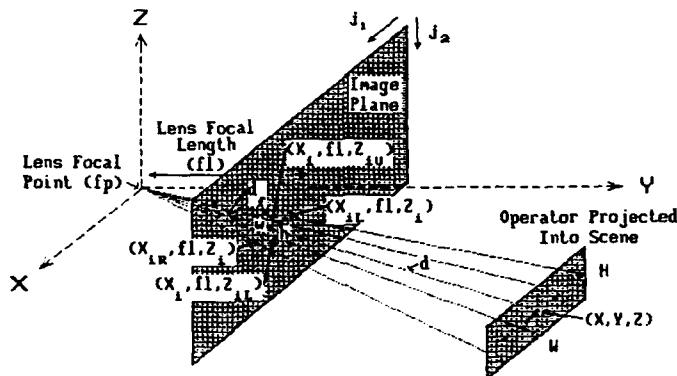
Figure 5: Morphological window operator, W_h 

Figure 6: 3-D Scene to 2-D Image Projection

The autonomous vehicle application requires that only regions of two textures be identified; the low texture pathway region, and high texture background. Detection of the pathway within the image subfield can now be viewed as an image filtering problem. The continuous line image across the low texture pathway region can be treated as the object of interest, whereas the broken portions of the line across the background regions appear as noise. Pathway detection amounts to filtering the noise while accurately preserving the edges of the object. A ranked order filter is chosen because it is particularly effective in a situation where the noise consists of strong impulses, such as the broken line clusters encountered in the background regions, and where the image characteristic to be preserved is object edge sharpness [9,10].

A morphological window operator is defined where M is the width of the window in pixels, and is odd, and N is the height of the window in image subfield lines (see Fig. 5). These parameters are established in scene dimensions based on vehicle and scene characteristics. For example, the height of the operator projected into the scene should equal the acceptable texture height in the scene, which should be something less than the vehicle ground clearance. The projected operator width should be greater than the widest unacceptable continuous plane in the scene (i.e. width of a large tree, leaf, or stone).

The relationship of operator parameters, as defined in the scene, to their projection into the image plane is illustrated in Fig. 6. The values of these projected parameters in the image plane are functions of a laser plane and location in the image; however, all of these parameters can be calculated in advance and stored in a look-up table. For each laser plane, it is possible to determine in advance, the required operator sizes at every pixel in the image. Fig. 6 depicts the process used to determine operator size at a particular pixel, for an arbitrary laser plane. The scene coordinates of the point,

(x, y, z) , where the ray extending from the focal point through the image plane intersects the laser plane, can be obtained using Eqns. 1 through 6. The actual coordinates of the pixel of interest, (x_i, f_l, z_i) , are determined utilizing the results of a camera calibration. Definition of the acceptable texture height, H , and widest unacceptable plane, having a width of W , occurs in a plane, containing the scene point, (x, y, z) , that is parallel to the xz -plane. With the operator parameters defined in the scene, calculation of the operator size in the image plane is accomplished using the equations below.

$$w = W \sqrt{\frac{x_i^2 + f_l^2 + z_i^2}{x^2 + y^2 + z^2}} \quad (8)$$

$$h = H \sqrt{\frac{x_i^2 + f_l^2 + z_i^2}{x^2 + y^2 + z^2}} \quad (9)$$

Subsequently, the height and width of the image plane in actual units, h_i and w_i , and in pixels, J_1 and J_2 , are utilized in determining operator size in pixels, as given in the following equations. Note, however, that the resulting operator height and width, N and M respectively, are quantized to whole pixel values, thus, a particular size operator is often applicable to large regions of the image.

$$M = \frac{w}{w_i} J_1 \quad (10)$$

$$N = \frac{h}{h_i} J_2 \quad (11)$$

The window operator is translated throughout the image subfield. The location of the window (i_0, l_0) , where i_0 is the horizontal window center and l_0 is the top line, is described by:

$$W_h [t_h(i_1, i_2); i_0; l_0] = \left\{ t_h^l(i_1) : i_0 - \frac{M}{2} \leq i_1 \leq i_0 + \frac{M}{2}; l = l_0, l_0 + N - 1 \right\} \quad (12)$$

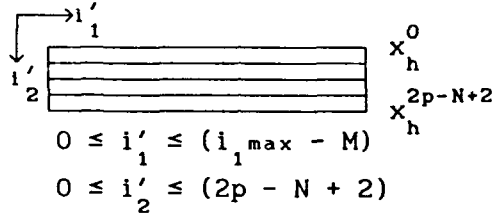
A ranked order filtering operation is performed at each window location, resulting in a mapping of each image subfield, $t_h(i_1, i_2)$, to a filter output subfield, $x_h(i'_1, i'_2)$. Each pixel in the filter output subfield corresponds to the filter output:

$$x_h(i'_1, i'_2) = \text{Rank}_T [t_h(i_1, i_2) \in W_h[i_0, l_0]] \quad (13)$$

where the Rank_T operation consists of ordering the pixel values within the filter window from 1 to 0 and choosing the pixel value at the rank corresponding to the threshold T . Thus, if T or more pixel values are 1, the filter output is 1. The value of T is related directly to the operator width such that a near continuous line across the operator is detected.

The filter output subfield (Fig. 7) is then compressed vertically into a single line, $y_h(i'_1)$, corresponding to the image field spatial location of the Hough transform extracted line, $t_h(i_1)$, where:

$$y_h(i'_1) = \text{Max}[x_h^0(i'_1), x_h^1(i'_1), \dots, x_h^{2p-N+2}(i'_1)] \quad (14)$$

Figure 7: Filter output subfield, $x_h(i'_1, i'_2)$

Although the ranked filter preserves the pathway edges, or ends of the pathway lines, the filtering operation erodes $M/2$ pixels from each end of each line and both sides of any breaks in the lines. Restoration of eroded pixels is accomplished by morphological dilation, which is a transformation that combines two sets using vector addition of set elements. If A and B are sets in Euclidean N -space (E^N), the dilation of A by B is denoted by $A \oplus B$ and defined by [11]:

$$A \oplus B = \{c \in E^N \mid c = a + b \text{ for some } a \in A \text{ and } b \in B\} \quad (15)$$

The vertical compression of the filter output subfield into a single line (y_h), results in the dilation needing to be performed only in the single horizontal spatial dimension, therefore, $N = 1$ in the above definition. Each of the resulting filtered, compressed lines is then dilated with an M by 1 structuring element, B :

$$y'_h(i'_1) = y_h(i'_1) \oplus B \quad (16)$$

The dilated line segments, $y'_h(i'_1)$, overlay the region of low texture pathway crossed by the structured light horizontal grid line, and the left and right end points, $y'_h(i_{1l})$ and $y'_h(i_{1r})$, of each line segment correspond to the left and right boundaries of the pathway in image subfield coordinates.

These end points, $y'_h(i_{1l})$ and $y'_h(i_{1r})$, are then mapped to their corresponding image coordinates, $a_{hl}(j_{1l}, j_{2l})$ and $a_{hr}(j_{1r}, j_{2r})$, using the transformations:

$$j_1 = i_1 \text{ and } j_2 = a_h i_1 + b_h \quad (17)$$

The points, $a_{hl}(j_{1l}, j_{2l})$ and $a_{hr}(j_{1r}, j_{2r})$, constitute a pathway boundary point pair which, when combined with all other such pairs ($h = 1$ to h_{MAX}), defines the pathway location in the image field quantized corresponding to the number of horizontal structured light grid lines.

In practice, pixels are not extracted from the image field and placed in a new location and orientation. Instead, the window operator is structured with slope a_h and translated through the image field, $t(j_1, j_2)$, along lines parallel to the line $j_2 = a_h j_1 + b_h$. The output of this process is a boundary point pair detected at each horizontal laser line image location. Connecting all of the left pathway boundary points with straight line segments, and all of the right points in the same manner, yields an image of the estimate of the pathway boundaries.

C. 3-D Point Extraction

Image segmentation utilizes only part of the information obtainable from the structured light. The discrete horizontal grid lines correspond to known vertical laser projection angles. As a result

of the segmentation processing, the approximate vertical image location of each horizontal laser line corresponds to the vertical position of the center line, t_h^0 , of the image subfield, $t_h(i_1, i_2)$, in which it is contained. Also, the 3-D coordinates of only the points within the pathway boundaries are of interest, therefore an arbitrary number of image points can be chosen lying on the image subfield center line between the previously identified left and right boundary points, $a_{hl}(i_{1l}, i_{2l})$ and $a_{hr}(i_{1r}, i_{2r})$, in each subfield. All points in one image subfield will have the same vertical laser projection angle.

As in the segmentation procedure described above, the chosen subfield points are mapped to their corresponding points in the image field using the transformations shown in Eqns. 17. By using time sequential images of single laser lines, the grid line that corresponds to each chosen point is identified, such that the vertical laser projection angle of each chosen point can be determined. The image coordinates, (j_1, j_2) , of each chosen point can then be mapped to camera image plane coordinates using the camera model and calibration described in Section 2B. The laser angle and camera image coordinates constitute all the information required to utilize Eqn. 6 to calculate the 3-D scene coordinates corresponding to each detected point.

Once again, in practice, pixels are not extracted from the image field and placed in a new location and orientation described as a subfield. Instead, point selection is accomplished along each line $j_2 = a_h j_1 + b_h$ from the output of the Hough transform, and need only occur between the previously identified left boundary point, $b_{hl}(j_{1l}, j_{2l})$, and the identified right boundary point $b_{hr}(j_{1r}, j_{2r})$.

4. Experimental Results

A. Scene Segmentation

The experimental "autonomous vehicle", utilized to test the previously described processing procedures, consisted of a tripod mounted video camera and a separately mounted portable holographic laser scanner. Numerous laboratory and outdoor scenes were imaged and processed during design and testing of the described processing. These images and the corresponding results are detailed in [12]. Analysis of one particular full scale outdoor pathway scene yielded the images shown in Figs. 8-12. Fig. 8 is a captured video image of the uniformly lit scene containing the pathway extending away from the camera in the y direction. The low texture path in the center of the image is nearly the same color as its high texture surroundings and thus, when shown as a 2-D image, is difficult to detect with the human eye or by grey scale methods. Fig. 8 also shows the preprocessed, horizontal laser projected light grid overlaid on, and registered with, the scene image.

When the structured light horizontal grid was projected onto the scene, the laser lines became discontinuous in proportion to the 3-D texture in various regions of the scene. As shown in Fig. 8, the laser lines became highly discontinuous over the high texture background in the scene and remained fairly continuous throughout the lower texture pathway in the center of the scene. The globally thresholded binary image field, $t(j_1, j_2)$ shown in Fig. 9, was utilized for further processing (image polarity has been reversed for clarity).

To locate the laser lines in the image, the Hough transform was performed on the image of Fig. 9. Image subfields, shown in Fig. 10, were then extracted from the image field where the center line (t_{h_0}) of each subfield, $t_h(j_1, j_2)$, corresponded to a line detected by the Hough transform. Each subfield was five ($2p + 1$ with $p = 2$) image lines wide.

The next goal in the processing was to eliminate the apparent image noise associated with the high texture background while preserving the boundary information associated with the low texture pathway. This was accomplished by using a ranked filter with an M by N window operator and subfield that varied in size from image top to image bottom such that their sizes in the scene correspond roughly to sensing 5.6cm wide by 3.1cm high along the path segment. Scaling these dimensions up to a reasonable pathway width would correspond to a 10cm vertical acceptance and a horizontal rejection width of 18cm. Ten cm is small, but still within a range of acceptable vehicle ground clearance, and 18cm is a reasonable largest probable unacceptable planar surface such as a flat spot on a tree or a flat rock.

A rank threshold equal to M was chosen because a horizontal contiguous line, with vertical deviations not exceeding N pixels, contained within the window would illuminate at least M pixels. The seven subfields, $t_h(j_1, j_2)$, had axes aligned with the subfield's orientation yielding the new coordinates, $t_h(i_1, i_2)$ where:

$$i_1 = j_1 \text{ and } i_2 = a_h j_1 + b_h + l, -p \leq l \leq p. \quad (18)$$

In theory, all operations were performed in the (i_1, i_2) plane, however in practice, all steps were performed in the (j_1, j_2) plane. The filter's output, $x(i'_1, i'_2)$, could then be described by Eqn. 13, and subsequently compressed vertically as described by $y_h(i'_1)$ in Eqn. 14. The output of such a filter in the (j_1, j_2) plane then corresponded to white pixels located along lines that had position and orientation of the lines previously derived from the Hough transform. Restoration of eroded pixels was achieved by morphological dilation of all $y_h(i'_1)$ with an M by 1 structuring element, B , as defined by Eqn. 15. The result of this dilation is shown in Fig. 11. Using the slope and intercept information obtained from the Hough transform, the left and right boundaries of the path at each laser line were located by searching along each Hough line for the farthest right and left white pixels. All boundary points on the left side of the path were then connected, and all boundary points on the right side of the path were connected, yielding the path outline shown in Fig. 12 overlaid on the original scene image.

B. Pathway Definition

Knowledge of the relative positions of all visual system components is a requirement for implementing a stereoscopic visual system. To achieve this, a theodolite was utilized for measuring the horizontal and vertical angles of various points throughout the neighborhood of the experimental "autonomous vehicle." The additional measurement of distances, from theodolite to each point, provided a set of spherical coordinates for each surveyed point. Visual system descriptors included in the survey were the focal points of the laser scanner and camera lens. Eight points along the perimeter of the pathway were also surveyed. The planar fit algorithm, described in Section 2 and the Appendix, was applied to these points, thus providing a basis of comparison for the resulting 3-D points and planar fit derived by the vision system.

The video camera used in all experiments was calibrated according to the procedure previously explained. The calibration information was used to calculate the vertical and horizontal image plane displacements corresponding to each chosen point. To derive the laser angles of projection, a look-up table had been constructed as described in Section 2. This table contained the vertical laser angles of projection for every horizontal line in the projected grid. Thus, determination of the particular

scan line that contained the point allowed reference to the look-up table to obtain the corresponding angular information.

After the camera image location and vertical laser projection angles of each chosen point had been derived, scene definition could proceed. The method described in Section 2 was used to calculate the 3-D coordinates of all the chosen points. Finally, the planar fit technique, described in Section 2 and the Appendix, was applied to the calculated points in the sample image, producing the equation of the plane of the sample pathway as derived by the vision system. The derived equation for this particular experimental pathway is given in Eqn. 19 for comparison with the planar fit of the surveyed data points of the pathway which is described by Eqn. 20.

$$0.08x + 0.27y - 0.96z - 199.36cm = 0 \quad (19)$$

$$0.01x + 0.23y - 0.97z - 197.94cm = 0 \quad (20)$$

5. Discussion

A. System Performance

Visual examination of Fig. 12 clearly indicates that the vision system described herein is capable of segmenting a scene in an unstructured environment into regions of background and likely pathway. The detected boundaries are well defined and, as determined by comparison with survey data, do coincide with the actual pathway boundaries. The output of the segmentation procedure consists of a minimal amount of data (seven pairs of points in this case) which results in a succinct pathway identification that will provide desirable input to a vehicular navigation system.

In addition, the consistency of Eqns. 19 and 20, demonstrates that this vision system is not only capable of discerning the pathway ahead of the vehicle, but also is capable of locating the pathway in 3-dimensions and defining it as a plane. Although the 4.7° angle between the normals of the two planes appears to be significant to vehicular navigation, it should be realized that the major contributing component is the difference in angles with respect to the x axis. This rotation is caused by the resolution in point location as projected across the width of the pathway. The average width of this particular pathway was $93cm$, which is substantially scaled down from an expected width of an acceptable vehicle pathway. With a $93cm$ base length, a 4.7° rotation can be caused by a point z component inaccuracy of as little as $6cm$. Such a vertical inaccuracy, when measured at a distance from laser to path center of $830cm$ in the y direction, can result from an angular error on the order of 0.4° . This angular error is beginning to approach laser beam divergences and the projection angle accuracies defined for this experiment. In addition, the $4cm$ vertical displacement between the derived and surveyed planes is insignificant in terms of vehicular navigation. In all, the pathway definition is considered correct within the experimental constraints, and it is expected to improve significantly as the pathway size is increased.

B. Computational Simplicity

The described segmentation and definition procedures utilize standard image processing methods, such as the Hough transform and image morphology. Very little unique processing is involved, and therefore, the implementation of this procedure is well defined and computationally simpler than more elaborate methods. This reduction in computational complexity should result in a substantial increase in speed over current image segmentation methods, and should also make it possible to implement the

unique algorithms in hardware for still greater processing speed, or on computers which could easily be integrated into an autonomous vehicle system.

Image morphology is beginning to see widespread application, and as a result, considerable effort is being devoted toward hardware implementations that considerably increase the speed of various morphological processes. Hardware implementation efforts are also being directed toward the Hough transform, and this could result in substantial reduction in processing requirements and increased computational speed.

C. Image Definition

The research described herein concentrated on a single planar fit to the region of interest in each scene. This was done for simplicity and because of general applicability to vehicular navigation, where it is likely that any acceptable pathway will be planar within some tolerance. In fact, this tolerance could be used as one condition for pathway acceptability. An extension to the planar fit method, described in the Appendix, can be utilized to determine when the region of interest in the scene contains bifurcations requiring a multi-planar or piece-wise planar fit. This extension could merely be a routine that decides when the eigen value associated with the best single planar fit, described in the Appendix, exceeds a predetermined threshold for "planar". An algorithm would then be invoked which would determine the image location of a likely bifurcation, and regions of the pathway for two or more separate planar fits. The information could also provide input to a higher level of processing which is capable of deciding the meaning of detected planar orientations and break points, such as a steep slope, cliff, or wall in the pathway. Although the above described higher level processing could be classed as artificial intelligence, and is beyond the scope of this research, it should be noted that the methods described in this paper can provide direct interface with such processing.

D. Obstacle Location and Identification

As described in Section 2, the image of structured light responds to the 3-D shape of the scene. The assumption was made during this research that there were no obstacles in the chosen pathway. However, an obstacle in the pathway is simply a distortion in the 3-D shape of the pathway, and is therefore sensed by all of the processing methods described. An additional step following the dilation of the filtered image would be required to examine each scan line image for discontinuities that could indicate the presence of an obstacle in the region that is about to be chosen as pathway. Pathway select versus obstacle alert criteria would be dictated by higher level processing and control systems.

An adaptive method may be applicable whereby initial processing identifies a discontinuity in pathway scan lines that may indicate the presence of an obstacle, and the laser scanner would then be aimed at that area to provide higher resolution scanning for identification. The laser scanner utilized during this research can be constructed with the dynamic capability to direct higher resolution scanning at selected portions of the original larger scan area. The scene definition methods described previously could then be applied to this higher resolution, small scan area to define the obstacle.

Appendix: Planar Fit

Suppose the points, (x_k, y_k, z_k) , for $k = 1, 2, \dots, n$, on a particular surface are available. Determination of the best fitting plane is achieved by minimizing the sum of the squares of the distances, from each point, to the plane. The plane that is sought is described as $ax + by + cz + d = 0$. The sum

to be minimized is thus

$$\sum_{k=1}^n \frac{(ax_k + by_k + cz_k + d)^2}{a^2 + b^2 + c^2}. \quad (21)$$

By requiring that the magnitude of the normal to the plane equal unity, $a^2 + b^2 + c^2 = 1$, the process is transformed into a constrained extremum problem. Therefore, the method of Lagrange multipliers can be applied resulting in the equation given below, where λ denotes a Lagrange multiplier.

$$\nabla \sum_{k=1}^n [(ax_k + by_k + cz_k + d)^2] = \lambda \nabla (1 - a^2 - b^2 - c^2) \quad (22)$$

Subsequent expansion and manipulation of Eqn. 22 yields a system of equations that can be expressed in matrix form as illustrated by Eqn. 23. Note that the terms x_a , y_a , and z_a denote average values of x , y , and z , respectively (ie. $x_a = \frac{\sum x_k}{n}$). Also, the multiplier, λ , in the matrix equation is an eigen value.

$$\left[\begin{bmatrix} \sum x_k^2 & \sum x_k y_k & \sum x_k z_k \\ \sum x_k y_k & \sum y_k^2 & \sum y_k z_k \\ \sum x_k z_k & \sum y_k z_k & \sum z_k^2 \end{bmatrix} - n \begin{bmatrix} x_a^2 & x_a y_a & x_a z_a \\ x_a y_a & y_a^2 & y_a z_a \\ x_a z_a & y_a z_a & z_a^2 \end{bmatrix} \right] \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \lambda \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (23)$$

Solution to this matrix equation consists of three eigen values along with their corresponding eigen vectors. The normal, $N = \langle a, b, c \rangle$, to the plane of best fit is the eigen vector that corresponds to the smallest eigen value. Then, the parameter, d , is readily calculated by utilizing the resulting d component of the gradient relationship expressed in Eqn. 22.

Acknowledgements

The author wishes to acknowledge the contributions of Mr. Richard Marvel, Dr. Gonzalo Arce, Dr. Charles Ih, and Mr. David Fertell throughout various parts of the research and experimentation. Also, Dr. Charles Boncelet proposed the method for computation of the planar fit.

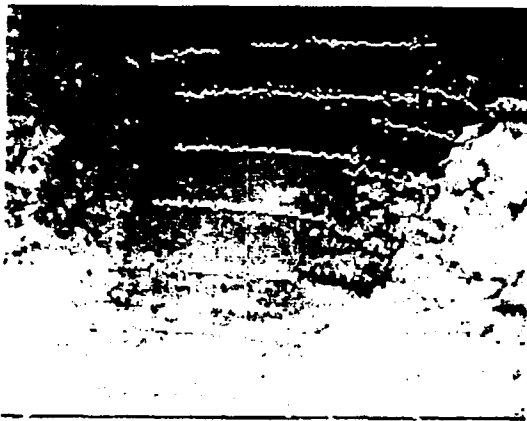


Figure 8: Video image of the scene with overlaid structured light

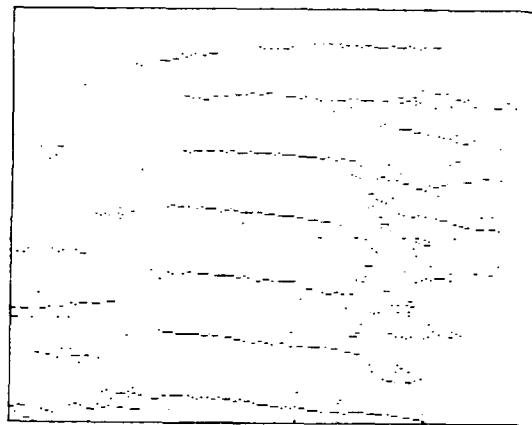


Figure 9: Binary image of structured light

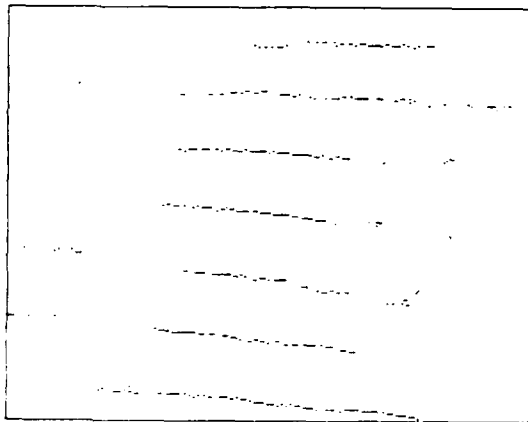


Figure 10: Image subfields extracted with Hough transform

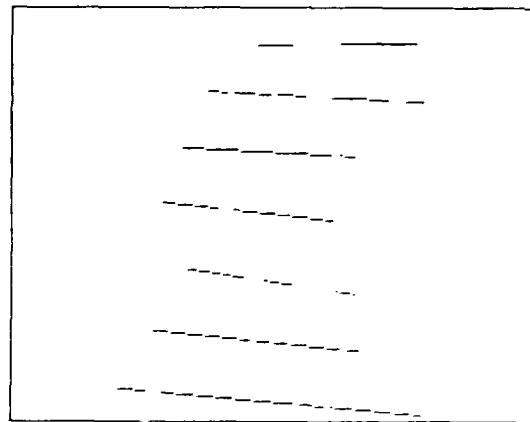


Figure 11: Centerline of each subfield following morphological erosion and dilation



Figure 12: Detected pathway boundaries overlaid on original scene

References

- [1] J. LeMoigne, A. M. Waxman, "Projected Light Grids for Short Range Navigation of Autonomous Robots," *Computer Vision, Graphics, and Image Processing*, vol. 28, pp. 203-206, 1984.
- [2] A. M. Waxman, J. J. LeMoigne, L. S. Davis, B. Srinivasan, T. R. Kushner, E. Liang, and T. Siddalingaiah, "A Visual Navigation System for Autonomous Land Vehicles," *IEEE Journal of Robotics and Automation*, vol. 3, no. 2, pp. 124-141, April 1987.
- [3] R. M. Inigo, E. S. McVey, B. J. Berger, and M. J. Wirtz, "Machine Vision Applied to Vehicle Guidance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 820-826, Nov. 1984.
- [4] D. Nitzan, "Three-Dimensional Vision Structure for Robot Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 3, pp. 291-309, May 1988.
- [5] S. J. Gordon and W. P. Seering, "Real-Time Part Position Sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 3, pp. 374-386, May 1988.
- [6] P. J. Besl and R. C. Jain, "Segmentation Through Variable-Order Surface Fitting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 2, pp. 167-192, March 1988.
- [7] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd ed., Reading, MA: Addison-Wesley, pp. 130-134, 1987.
- [8] J. Serra, *Image Analysis and Mathematical Morphology*, New York: Academic Press, pp. 15-20, 1982.
- [9] P. Maragos and R. Schafer, "Morphological Filters - II: Their Relationship to Median, Order-Statistics, and Stack Filters," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1170-1184, Aug. 1987.
- [10] T. Nides and N. C. Gallagher, "Median Filters: Some Modifications and their Properties," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, Oct. 1982.
- [11] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image Analysis Using Mathematical Morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 532-550, July 1987.
- [12] C. D. Brown, *Scene Segmentation and Definition for Autonomous Robotic Navigation Using Structured Light Processing*, Doctoral Dissertation - University of Delaware, Dec. 1988.

Advanced Ground Propulsion System Phenomena

Dr. Walter Bryzik,* Mr. Ernest Schwarz, Mr. Eugene Danielson,
Mr. Michael Reid

U.S. Army Tank Automotive Command
Warren, MI. 48397-5000

INTRODUCTION

Advanced low heat rejection (LHR) diesel engines offer dramatic opportunities for quantum improvements within future Army ground vehicles. An early evidence of this opportunity is seen by viewing the new upgraded M109 Howitzer and Paladin which use LHR technology developed at TACOM.

Within the LHR engine concept, the diesel combustion chamber is insulated using high temperature materials on engine components such as pistons, cylinder head, valves, cylinder liner and exhaust ports (1,2). Additional power and improved efficiency derived from this concept occur because thermal energy normally lost to the cooling water and exhaust gas is potentially converted to useful power. By greatly reducing lost energy and essentially eliminating the need for a conventional cooling system, (i.e. all water is removed and strategic oil cooling employed) this engine system will dramatically improve overall performance and could potentially result in a 50% reduction in both weight and volume for the entire propulsion system (i.e., engine, cooling system, air intake system, exhaust system and fuel system including fuel quantities). Such an engine would have great positive impact particularly upon military tactical and combat vehicles. Elimination of the engine cooling system including cooling fans, radiators, hoses and shrouds would produce a quantum jump in reliability and maintainability. The engine would not be sensitive to conventional cooling damage and extreme environmental conditions. Specific weight reductions allow improved vehicle response, while specific volume reductions allow reduced armor cover requirements and improved survivability characteristics. Because of high temperature engine operation, quieter combustion and improved multifuel characteristics are possible. With this engine concept, the entire philosophy of combat vehicle design becomes far less restrictive as concerns regarding satisfactory locations of cooling grilles, passages, and associated equipment are eliminated. Cost of the engine is expected to be equal or less than its cooled counterpart since water pump, seals, hoses and costly water jackets would be eliminated.

In order to maximize this overall LHR engine technology, a number of scientific challenges require fundamental focus including: 1) high temperature combustion optimization and 2) structural integrity of high temperature and complex engine components. This paper will address significant aspects of these scientific challenges.

OBJECTIVES AND METHODOLOGY

The primary objective of the research is to investigate what combustion factors may be responsible for performance differences between conventional watercooled and low heat rejection (LHR) or limited cooled diesel engines. Combustion aspects of the LHR engine have not been adequately studied and investigations of the past have yielded results that are often conflicting. Summaries of these results and discussion of reasons behind why results may be conflicting are extensively elaborated on in (3, 4, and 5). Analysis of these and in-house efforts have revealed that combustion phenomena within LHR engines are very complex and subject to extreme parameter sensitivities. Of particular note, complex interacting mechanisms within the LHR class of engine were seen to be extremely sensitive to heat release phasing and its breakdown between premixed and diffusion combustion. As an example, it was found that small changes in surface thermal conductivity will either significantly enhance or degrade engine performance depending upon characteristics such as engine geometry, heat transfer and fluid flow/chemical kinetic phasing interactions. These sensitivities are not nearly as great in conventional water cooled engines; and thus, fundamental combustion heat release/heat transfer mechanisms have dramatically different controlling mechanisms for LHR engines. The relationship between premixed and diffusion burning in a LHR engine is particularly critical. It is the intent of this research to further study combustion in an LHR engine and compare it against that in a standard watercooled diesel engine thereby giving more in-depth insight into the physical mechanisms related to changes occurring in LHR engine heat release rates and corresponding changes in engine performance.

High speed in-cylinder pressure data was gathered for a LHR-V903, Cummins single research engine at the U.S. Army Tank-Automotive, Propulsion Systems Division facility. This engine is a single cylinder version of the V-8 engine currently powering the Bradley Fighting Vehicle and is representative of military high output diesel engines. Also this single cylinder engine has the capability to run as a fully cooled engine or as a limited cooled (head cooling only configuration) engine. In-cylinder pressure data was obtained with varied coolant, intake manifold and oil temperatures for insulated and noninsulated piston cases. In the case of insulated pistons, porous zirconia and seal coated zirconia coatings were examined to determine effects on combustion and performance. From the in-cylinder pressure data, heat release rates were calculated and evaluated for a matrix of engine operating parameters and design configurations.

Combustion and performance data generated through this effort will also be used to improve in-house diesel cycle simulation codes. The current technical approach in these simulation codes is to use correlated results from experimental data. Predictions of internal operating conditions and performance of watercooled diesel engines are very good, however, predictions for LHR engines are in question. Actual LHR heat release data will be used to improve burning rate subroutines within the simulation to more accurately predict engine performance particularly for the case of LHR engines. Fundamental understanding (with experimental verification) in these areas is essential if future engine performance optimization is to be achieved.

TEST ENGINE

A single cylinder research version of the Cummins VT-903 Bradley Fighting Vehicle engine was used in this overall effort. The engine, was made by cutting the block of a standard V-8 and leaving only two cylinders at the flywheel end. Only one cylinder is active while the other has a slipper piston which does no work but is used for balancing. The front end of the engine, the crank, cam, cylinder head and fuel

injection system were modified accordingly to match the single cylinder configuration. The VT-903 engine is currently used in the Bradley Fighting Vehicle, the M2, and is representative of high output diesels used for combat vehicle applications. Engine specifications are provided in Table 1. This engine is configured so that cooling of the head alone is possible thereby simulating low heat rejection or minimum cooled operation.

A ductile iron piston was used for the standard cooled engine configuration. Two different pistons were fabricated for IHR minimum cooled tests. The first of these was a 0.040 inch thick plasma sprayed zirconia coated ductile iron piston and the second, the same except that the surface was impregnated with chrome oxide which acted as a seal coat. Precautions were taken to ensure that the insulated test pistons had the same compression ratio as that of the baseline case which was 13.9. The test pistons are illustrated in Figure 1.

A synthetic, polyester lubricant, Stauffer SDL-1 was employed throughout all engine tests. This lubricant has been extensively used during Army and other IHR engine developments proving its adequacy for the temperature ranges encountered during these tests. Standard diesel fuel was used for all tests using a specification of VVF-800, DF-2.

Table 1 VT-903 SINGLE CYLINDER
RESEARCH ENGINE SPECIFICATION

TYPE	4 CYCLE DI DIESEL
BORE	5.6 IN
STROKE	4.75 IN
DISPLACEMENT	112.88 CU IN
RATED SPEED	2600 RPM
COMPRESSION RATIO	13.9
RATED BMEP	170 PSI
NO. OF VALVES	4
INJECTORS	CUMMINS PT

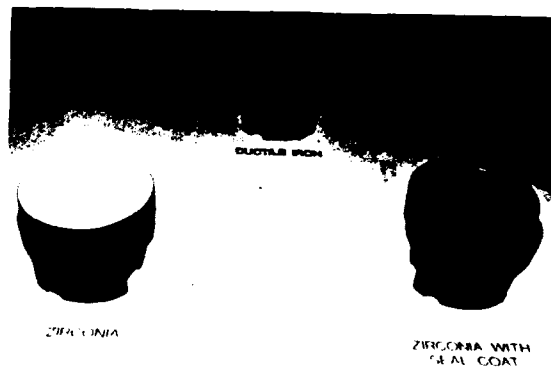


Fig. 1 Test Pistons

TEST INSTALLATION

The V-903 research engine was installed in Cell 7, at the U.S. Army TACOM, Propulsion Systems Division, Building 212 test facility. The engine was coupled to a Pohl eddy current dynamometer with a 125 HP absorption capability. This dynamometer set-up also had an induction motoring section (30 HP) for engine starting and friction testing. Figure 2 illustrates the engine installation along with accompanying support equipment. A Digalog Corp. system was used for dynamometer speed and load control and a load cell used for torque measurement.

Engine test cell support systems for air, fuel, cooling and oil systems were installed for maximum flow and temperature control flexibility. Thus, the engine operating conditions could be governed to accurately simulate those of a turbocharged, aftercooled engine within a combat vehicle application. This control of engine subsystems also provided the capability to perform parametric analyses to determine operating effects on performance and heat release. External air was used to simulate turbocharged conditions. Maximum capability for charge air pressure with this system is 90 PSI. Intake air was heated with a 120K btu/hr triple pass heater system so that temperatures could be equivalent to aftercooled air temperatures. Intake pressure was controlled with a

regulator and exhaust back pressure maintained with an electrically controlled butterfly valve, thus permitting an intake to exhaust manifold pressure ratio of one to be maintained throughout the testing. This condition is representative of turbocharged diesel engine applications. Cooling and lubrication systems were designed and installed which permitted accurate control of temperatures via pneumatic controllers. Flow rates could be controlled with pump bypass circuits. Coolant and lubricant circuits used shell and tube heat exchangers and steam heating was employed in order to facilitate bringing these fluids up to temperature without relying on engine heat. A Cummins PT injector was used in the research engine. Because the engine had no drive for an engine mounted PT pump, a variable speed fuel pump was employed. The fuelling circuit was equipped with a small needle valve bypassing the pump output allowing fine pressure adjustment. Fuel rate is controlled in a PT system via rail pressure control. Course fuelling was accomplished via adjustment of the pump speed.

Low speed data was acquired with a Fluke system. The engine was fully instrumented for temperature, pressure, speed, load and flow data. Fifty channels were employed for low speed data acquisition. Air, coolant, and lubricant flows were measured with turbine flowmeters. An Elwood system was used for fuel consumption measurement. High speed combustion data was acquired using a PEI Inc., PC based data acquisition hardware and analyzed using PEI Inc. ECA 892 matched software. Heat release analysis was also conducted offline using a modified Krieger Borman model developed at the University of Wisconsin with TACOM funding. The high speed system was used to record crank-angle resolved data to include cylinder pressure and fuel injector link load for the injection timing event. This system has four analog input channels with a sampling rate of 125K samples per second. All high speed data was taken over 100 cycles for each steady state operation point and ensemble averaged in order to provide a basis for statistically meaningful analysis. A Kistler 7061 piezo-electric transducer was employed for cylinder pressure measurement and BEI precision optical encoder with .2 CA degree resolution for marking and timing.

TEST OPERATING CONDITIONS

A test matrix for the ductile and insulated piston cases is given in Table 2. The engine was run at 1600, 2100 and 2600 RPM at 50% and 100% load levels for each speed. Torque levels for the baseline correspond to equivalent values of the V-8 engine at a 500 HP rating. Coolant, lubricant and intake manifold temperatures were varied to determine their effect on engine overall performance and heat release. Two insulative piston cases were originally planned to be run in the head cooled only configuration, however, only one insulated piston case was run and reported in this paper. Henceforth, this set of runs will be referred to as the LHR case.

Engine airflow and air to fuel ratio operating conditions are illustrated in Figure 3 for full load watercooled baseline tests. For all subsequent LHR tests the fuel rate was held constant and the airflow rate adjusted via intake manifold pressure to match the baseline case air to fuel ratios. Thus torque is made the dependent variable. This same method was used for the part load cases, although higher air to fuel ratios were run at these conditions which is representative of standard direct injected diesel engines. An intake to exhaust manifold pressure ratio of one was maintained for most test cases. Where this condition could not be achieved, baseline and LHR case ratios were made equivalent and thus noted.

Fig. 2 Test Installation

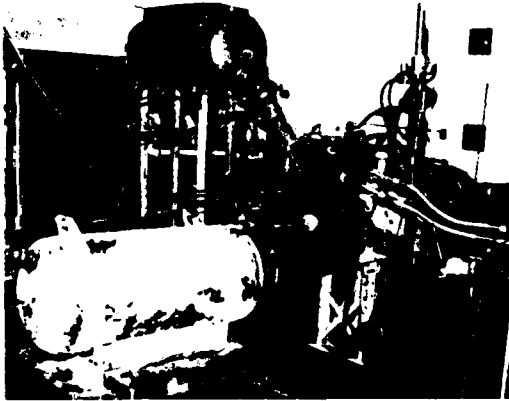


Fig. 3 AIRFLOW & A/F VS RPM
FULL LOAD

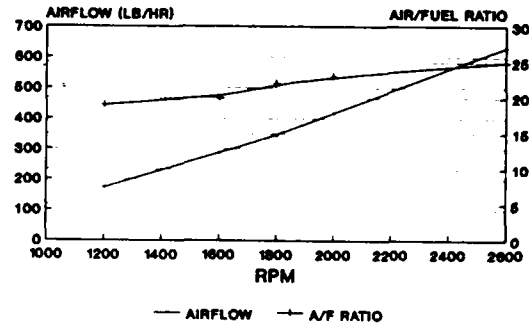


Table 2

TEST MATRIX

BASELINE DUCTILE IRON (A)

SPEED (RPM)	LOAD (%)	COOLANT TEMP (°F)	CHARGE AIR TEMP (°F)	OIL TEMP (°F)
1600	100	205	225	250
1600	100	205	140	250
1600	50	205	225	250
2100	100	205	225	250
2100	100	205	140	250
2100	50	205	225	250
2600	100	205	225	250
2600	100	205	140	250
2600	50	205	225	250

LHR II - INSULATED PISTON (ZIRCONIA)

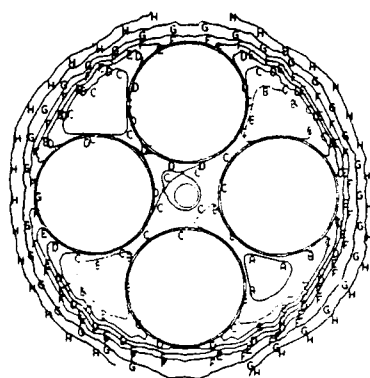
SPEED (RPM)	LOAD (%)	COOLANT TEMP (°F)	CHARGE AIR TEMP (°F)	OIL TEMP (°F)
1600	100	205	225	250
1600	100	250	225	250
1600	100	250	140	250
1600	100	300	225	300
1600	100	300	140	300
1600	50	250	225	250
2100	100	205	225	250
2100	100	250	225	250
2100	100	250	140	250
2100	100	300	225	300
2100	100	300	140	300
2100	50	250	225	250
2600	100	205	225	250
2600	100	250	225	250
2600	100	250	140	250
2600	100	300	225	300
2600	100	300	140	300
2600	50	250	225	250

THERMAL STATE OF THE TEST ENGINE

Estimates of temperatures across the firedeck were made using a thermal finite element analysis. A large scale, high resolution mesh with over 50,000 nodes was developed to model the firedeck section of the head. Figs. 4 and 5 present the isotherms across the firedeck for both the fully cooled and the LHR conditions. These temperatures are for a full power condition. Figure 6 shows temperature variation between valve bridge regions within the firedeck.

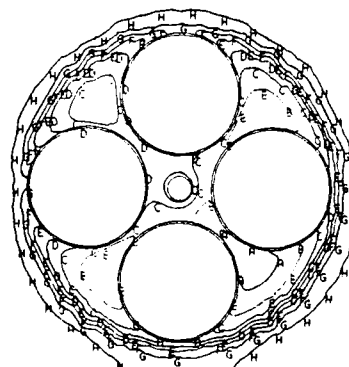
The estimated fully cooled peak temperatures agree well with those measured on another V-903 single cylinder engine (6) previously tested. On the basis of this agreement, it is assumed that the LHR isotherms are quite representative of actual temperatures. This means that the LHR head is operating at temperatures ranging from about 150°F to 200°F hotter than the fully cooled engine's head. At these elevated temperatures material selection could be a critical issue since the yield strength of grey iron begins to rapidly deteriorate above approximately 600°F. It would not be difficult to envision a potential yield condition in the partially cooled head.

The elevated temperatures are also seen in the liner which has no jacket cooling. Figure 7 presents a comparison of temperatures from prior engine testing (6) in the cylinder liner at the top ring reversal location. Here the LHR liner is operating at temperatures about 200°F to 250°F higher (depending on engine speed) than the liner in the fully cooled engine. Again, material selection may need to be carefully addressed in order to obtain long engine life particularly where lower heat rejection levels are sought.



DEGREES F

A = 616
B = 568
C = 520
D = 472
E = 424
F = 376
G = 328
H = 280



DEGREES F

A = 813
B = 744
C = 675
D = 607
E = 538
F = 469
G = 401
H = 332

Fig. 4 VT903 Fully Cooled Head Temperatures

Fig. 5 LHR903 Minimum Cooled Head Temperatures

Fig. 6 MAXIMUM TEMPERATURE ON FIREDECK
BRIDGE AREAS

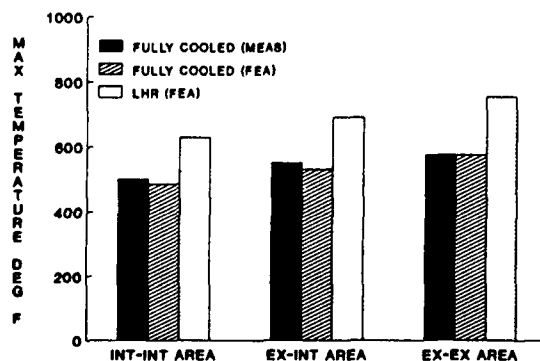
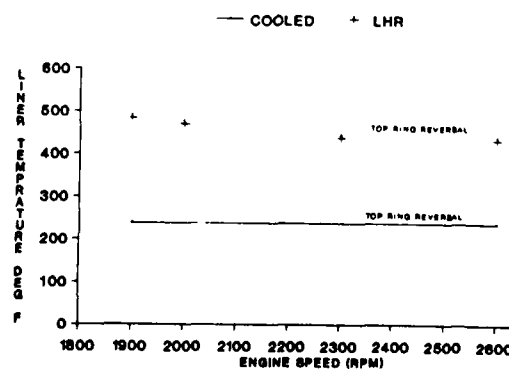


Fig. 7 FULLY COOLED VS LHR
LINER TEMPERATURE



The coated ductile iron piston thermal data reported for the partially cooled configuration in (6), indicates a maximum piston cap temperature of 1394°F, a substrate metal temperature of 750°F, top ring groove temperature of 620°F and maximum pin bore temperature of 320°F. Operating conditions in the current test program were similar to those reported in (6). Piston temperatures would be expected to be approximately 200°F cooler for the fully cooled configuration. This assumption is based on the general range of temperature differentials already established for the firedeck and liner. As with these other components, the piston material selection and/or coatings need to be carefully addressed.

Finally, the difference in specific heat rejection to coolant between the engines should be mentioned. The fully cooled engine in a multicylinder configuration has a specific heat rejection of 24.5 btu/hp-min while the LHR engine had measured values (6) in the range of 17.0 to 19.0 btu/hp-min dependent upon oil sump, coolant and charge air temperatures. This is considered a substantial decrease in heat rejection and allows for smaller cooling systems or less cooling fan parasitic losses.

COMBUSTION RESULTS

Pressure time histories, as well as injection timing events were recorded for all test points. From these histories, heat release rates were analytically determined. Pressure and heat release information are used to analyze the combustion process providing insight into the overall performance of the engine. Changes in combustion between the standard water cooled configuration and minimum cooled, ceramic coated piston configuration were readily apparent upon analysis of these diagrams.

Heat release diagrams for 1600, 2100 and 2600 rpm are provided in Figures 8 through 13. Within each particular figure, baseline and LHR cases are compared. All data for the LHR is at full load, 300°F coolant and 300°F oil temperatures. The first diagram in each rpm sequence for the LHR case is for a charge air temperature of 225°F and the second 140° F.

These heat release profiles are of the shape characteristic typically seen for quiescent or limited swirl direct injected diesel engines. Moving across these diagrams from left to right, one first encounters the premixed burning peak and then the diffusion burning peak. The diffusion burn portion for all these curves is the predominant mode of combustion. Analyzing these diagrams reveals some general trends given as follows: 1) the premixed burning peak decreases with increasing engine speed for a constant intake manifold temperature, 2) the premixed burning peak and burning portion is always greater at a given speed when intake manifold temperature is lower, 3) for the ceramic minimum cooled case, the premixed burning peak is always less than the conventional fully cooled engine with the difference being more at the low speed and very small at the high speed.

Further analysis of the heat release diagrams indicates a difference in beginning of ignition between the baseline and LHR case with the LHR ignition beginning one to three crank angle degrees earlier than in the conventional engine's case. The static injection timing for the Cummins PT injection system was fixed for all test runs. Initiation of injection with this injection system is a function of static timing and amount of fuel in the injector. The same amount of fuel was injected in the comparative cases. It has been noted on previous TACOM LHR development program efforts that the engines overhead may become so hot that the injector link may thermally grow causing an advance in injection timing. In this effort the injector link was strain gauged to determine approximate initiation of injection. The data revealed no apparent change in start of injection, therefore, the ignition phase shift

Fig. 8 HEAT RELEASE COMPARISON

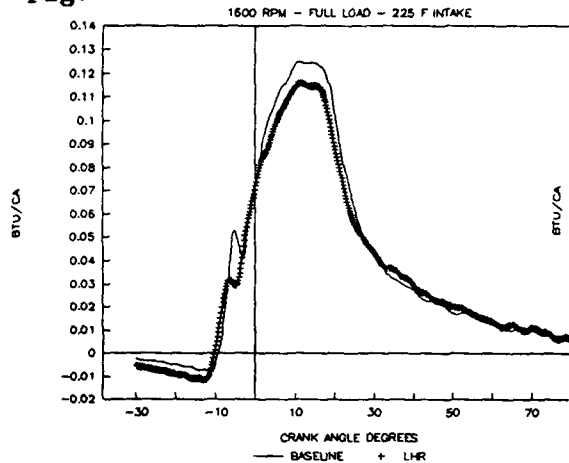


Fig. 9 HEAT RELEASE COMPARISON

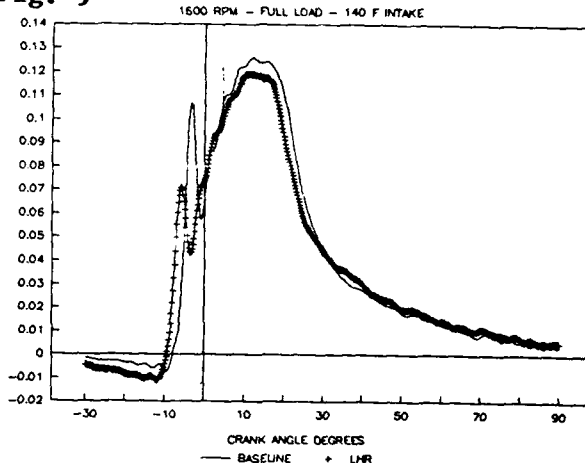


Fig. 10 HEAT RELEASE COMPARISON

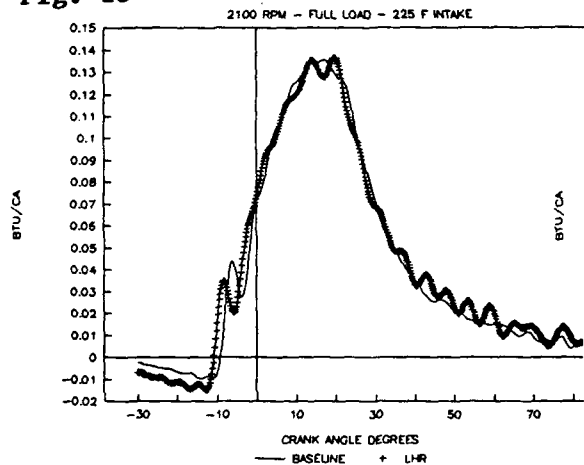


Fig. 11 HEAT RELEASE COMPARISON

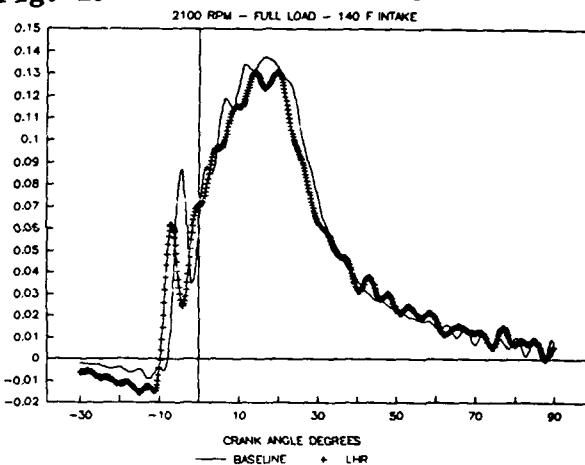


Fig. 12 HEAT RELEASE COMPARISON

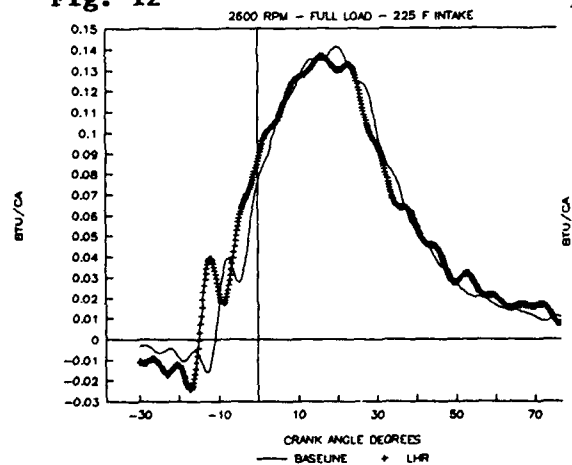
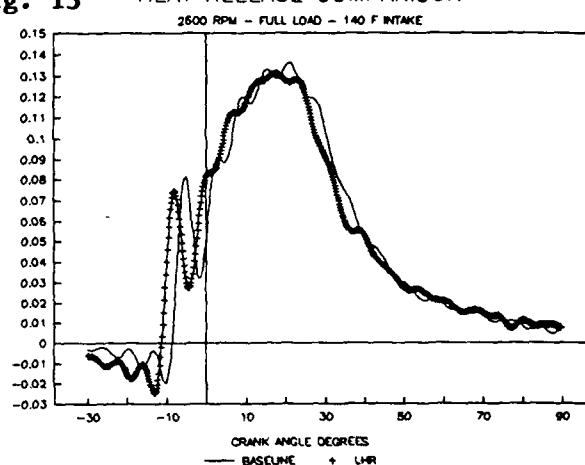


Fig. 13 HEAT RELEASE COMPARISON



displayed appears to be attributed to a change in ignition delay rather than any dynamic injection change. This makes sense as the higher temperatures encountered at time of injection would lead to a shorter ignition delay period and would also lead to a decreased premix fraction as was indicated in virtually all the heat release diagrams for the full load conditions.

Half load heat release comparisons were made for 1600, 2100 and 2600 rpm. Coolant temperature was 250°F, oil temperature was 250°F and charge air temperature 225°F for the LHR. Trend comparisons are not as dramatic as for the full load cases, however, they do tend to follow in that the premixed differences are more apparent at the lower speeds with less premixed burning for the LHR case. Burning duration was longer for the LHR at all speeds.

Heat release summary data for full load and half load is given in Tables 3 and 4. Included in these summaries are the peak cylinder pressures, the peak premixed heat release rate, peak diffusion heat release rate (peak for overall cycle) and the heat release duration. The heat release duration in crank angle degrees is calculated as the portion of the cycle from where .5% to .95% of the cumulative net heat release occurred. This summary data provides insight into how engine performance is affected by these parameters. As discussed previously, the premixed burning fraction was generally less for the LHR cases except at 2600 rpm full and half load for the hotter 225°F intake manifold temperature conditions in which instances they were equal. In all cases the heat release duration was greater for the LHR cases. This is clearly illustrated in Figure 14. The heat release duration increased as a

Table 3 HEAT RELEASE COMPARISONS - FULL LOAD
300 F COOLANT, 300 F OIL

CASE	SPEED (RPM)	INTAKE MANIFOLD TEMP (°F)	PEAK CYL PRESSURE (PSI)	PEAK PREMIXED NET HR RATE (BTU/CA)	PEAK DIFFUSION NET HR RATE (BTU/CA)	HEAT RELEASE DURATION (CA)
CERAMIC	2600	225	1657	0.040	0.137	69.4
BASELINE	2600	225	1577	0.040	0.141	62.2
CERAMIC	2600	140	1441	0.073	0.131	70.4
BASELINE	2600	140	1457	0.081	0.136	66.6
CERAMIC	2100	225	1303	0.034	0.135	65.4
BASELINE	2100	225	1310	0.043	0.135	61.2
CERAMIC	2100	140	1241	0.061	0.131	67.2
BASELINE	2100	140	1269	0.087	0.137	64.0
CERAMIC	1600	225	1045	0.032	0.115	65.8
BASELINE	1600	225	1110	0.053	0.125	64.0
CERAMIC	1600	140	1075	0.072	0.119	68.2
BASELINE	1600	140	1094	0.106	0.126	63.2

Table 4 HEAT RELEASE COMPARISONS - HALF LOAD
250 F COOLANT, 250 F OIL

CASE	SPEED (RPM)	INTAKE MANIFOLD TEMP (°F)	PEAK CYL PRESSURE (PSI)	PEAK PREMIXED NET HR RATE (BTU/CA)	PEAK DIFFUSION NET HR RATE (BTU/CA)	HEAT RELEASE DURATION (CA)
CERAMIC	2600	225	875	0.034	0.101	61.4
BASELINE	2600	225	992	0.033	0.095	56.2
CERAMIC	2100	225	747	0.055	0.094	60.6
BASELINE	2100	225	796	0.057	0.092	57.8
CERAMIC	1600	225	670	0.048	0.088	59.2
BASELINE	1600	225	681	0.056	0.095	52.8

function of speed for these full load runs. The premixed burn and heat release duration data is in agreement with findings of Alkidas (7), Gatowski (8), Dickey (9), Morel (10), Assanis (3) and others who have conducted high temperature direct injected diesel engine experimentation.

Of extreme interest is peak cylinder pressure data as this plays a very important role in the fuel consumption numbers to be presented later. A full load peak cylinder pressure comparison is displayed in Figure 15 for constant 225°F manifold temperature. From Figure 15 and Table 3, cylinder pressure data indicates: 1) at 1600 rpm, LHR pressures are lower than the baseline, 2) at 2100 rpm, LHR pressures are about the same as the baseline, and 3) at 2600 rpm, LHR pressures are approximately equal at 140 F intake manifold temperatures but greater than the baseline at the 225°F intake. Part load peak cylinder pressures as depicted in Table 4 were lower at all engine speeds for the LHR case.

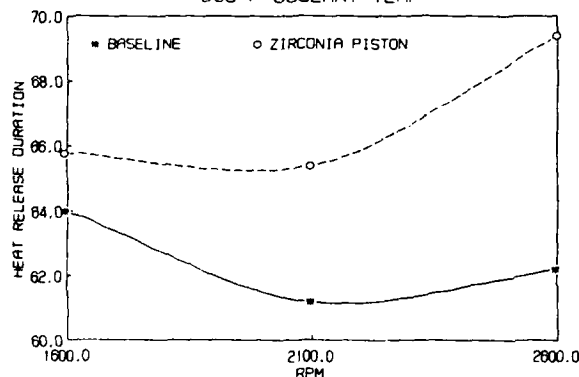
PERFORMANCE RESULTS

LHR full and part load engine performance data summaries are presented in Tables 5 and 6 respectively. Performance parameters include net indicated mean effective pressures and net indicated fuel consumption. These net indicated values are summations of the working and pumping loop mean effective pressures and fuel consumptions as calculated from the cycle pressure volume diagrams. Other relevant performance data includes intake pressures from which volumetric efficiencies can be calculated and the engine's exhaust temperatures.

As seen from the full load performance data, the LHR engine suffered a loss in volumetric efficiency as extra boost was required to maintain the constant air fuel ratio of the baseline. This difference was less than 2 in. Hg for all instances. Other researchers had experienced more severe losses for engines of this type with uncooled heads (11) but volumetric efficiency (based on intake manifold temperatures) could be regained (10) as intake air temperatures are increased. This particular V-903, LHR engine did not experience much loss in volumetric efficiency as the heads were still being cooled. All exhaust temperatures were higher for the LHR case with the highest difference of 113°F being experienced at the 2600 rpm, 225°F intake manifold point. Once again other researchers have seen higher exhaust temperature differences (8) where less cooling and more in-cylinder surface insulation were employed.

HEAT RELEASE DURATION

Fig. 14 300 F OIL TEMP
300 F COOLANT TEMP



CYLINDER PRESSURE COMPARISON

Fig. 15 300 F OIL TEMP
300 F COOLANT TEMP

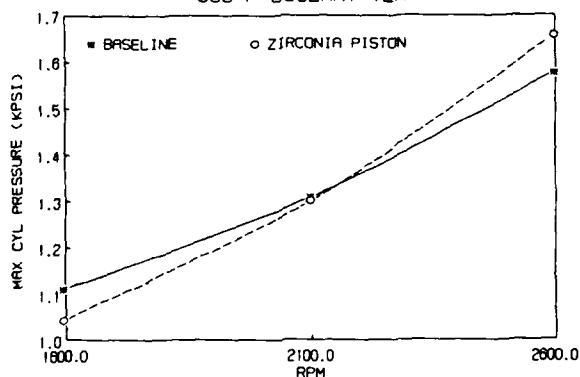


Table 5 PERFORMANCE COMPARISON - FULL LOAD
300 F COOLANT, 300 F OIL

CASE	SPEED (RPM)	INTAKE MANIFOLD TEMP (°F)	INTAKE MANIFOLD PRESS (in Hg)	EXHAUST MANIFOLD TEMP (°F)	NIMEP (psi)	NISFC (LB/NIHP-HR)
CERAMIC	2600	225	66.56	1377	208.2	0.343
BASELINE	2600	225	64.93	1264	202.5	0.353
CERAMIC	2600	140	55.97	1311	195.3	0.343
BASELINE	2600	140	55.36	1261	196.9	0.341
CERAMIC	2100	225	54.34	1312	188.8	0.327
BASELINE	2100	225	52.92	1237	188.2	0.329
CERAMIC	2100	140	48.24	1290	182.5	0.341
BASELINE	2100	140	46.61	1224	194.3	0.321
CERAMIC	1600	225	43.76	1312	162.0	0.353
BASELINE	1600	225	42.74	1255	171.7	0.334
CERAMIC	1600	140	39.69	1336	169.3	0.351
BASELINE	1600	140	38.88	1234	175.9	0.338

Table 6 PERFORMANCE COMPARISON - HALF LOAD
250 F COOLANT, 250 F OIL

CASE	SPEED (RPM)	INTAKE MANIFOLD TEMP (°F)	INTAKE MANIFOLD PRESS (in Hg)	EXHAUST MANIFOLD TEMP (°F)	NIMEP (psi)	NISFC (LB/NIHP-HR)
CERAMIC	2600	225	48.23	1036	105.6	0.362
BASELINE	2600	225	48.24	1019	105.9	0.362
CERAMIC	2100	225	38.88	1019	96.3	0.351
BASELINE	2100	225	38.88	998	107.3	0.315
CERAMIC	1600	225	36.23	921	87.9	0.348
BASELINE	1600	225	35.82	894	93.5	0.323

The fuel consumption comparison data is more clearly depicted in Figures 16 through 19. Figure 16 illustrates that the specific fuel consumption for the IHR engine was poorer at 1600 rpm, about equivalent to the baseline at 2100 rpm and better at 2600 rpm. At the time of the 2600 rpm run the author had some doubts about the data generated and consequently, re-ran that point at a later time but the same results were obtained. The trend of better fuel consumption with the higher in-cylinder temperatures and at the higher speeds is confirmed by examining Figures 17, 18 and 19 at the 2600 rpm points. Looking back at the combustion data presented earlier, the higher peak cylinder pressures encountered at the higher engine speeds and the less variation in the premixed burn for the IHR case would seem to substantiate why fuel consumption is better at the 2600 rpm rated speed condition. This was also borne out in Figure 19 where at 2600 rpm, the indicated specific fuel consumption improved as a function of increased head coolant temperature. Data presented in (5) done by Ricardo ITI and Purdue University on an engine with a similar type combustion chamber showed poorer ISFC at low speeds and better at high speeds, high loads for their IHR case as well. Data illustrated in Figure 18 which is a half load case once again shows improved performance at the 2600 rpm point, although here the IHR only matched the baseline fuel consumption. It should be noted, however, that this particular point is questionable as it was during this day's running that the piston experienced ceramic coating failure.

Fig. 16

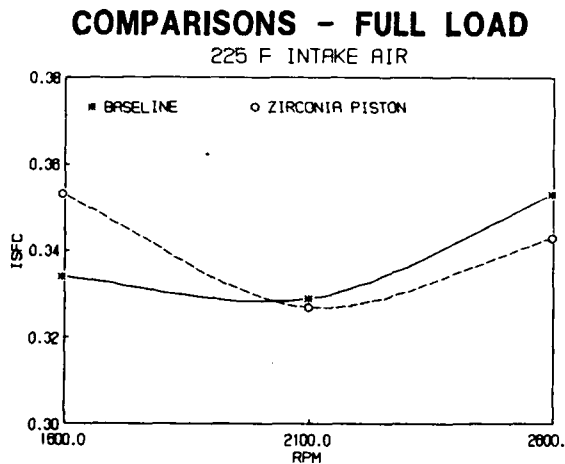


Fig. 17

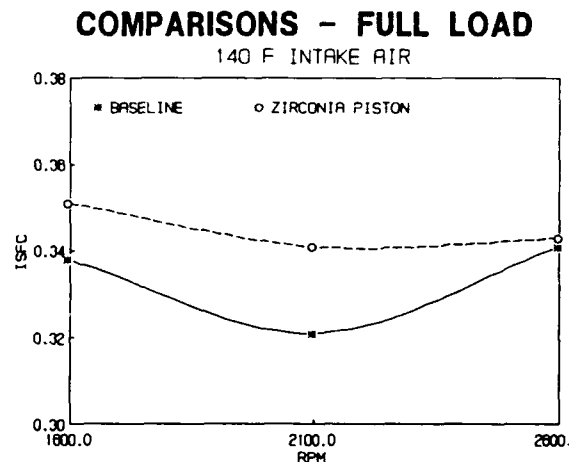


Fig. 18

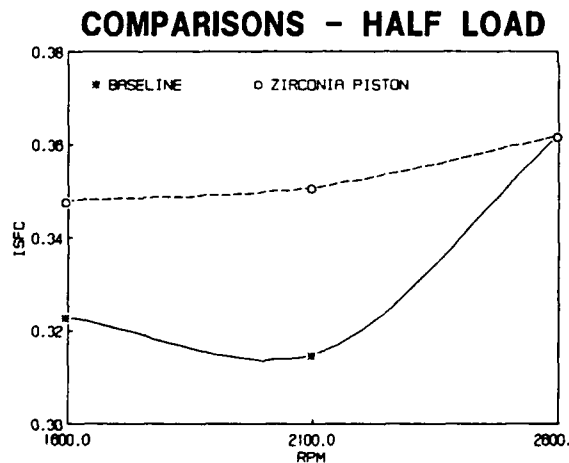
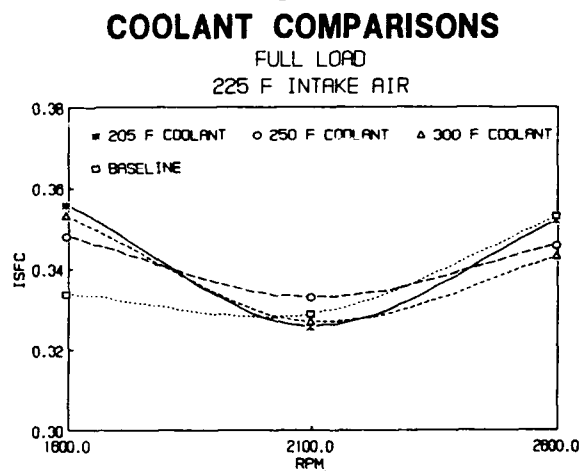


Fig. 19



The effect of blowby on heat release and performance is clearly expounded in (7) with severe blowby levels reported for IHR engines. Crankcase pressure was monitored throughout all the testing. Crankcase pressures were lower for all IHR test cases. Since this engine configuration had no liner cooling, piston/liner clearances were likely to have tightened up with resultant lower blowby particularly at the higher engine speeds. This, along with the ignition advance experienced could explain why cylinder pressures were higher for the IHR at the 2600 rpm point.

CERAMIC COATING DURABILITY

After each day's test runs, the cylinder pressure transducer was removed from the engine and the piston coating and bore examined to determine their condition because of concerns regarding the harsh operating environment of the engine. The plasma sprayed zirconia coating lasted approximately 60-hours before the coating began to spall from the ductile iron substrate. Upon engine teardown and piston examination, 50% of the coating was missing. All runs on that day were not reported within this write-up except for the 2600 half load point as previously noted.

Upon failure of the plasma sprayed zirconia piston, a chrome oxide seal coated zirconia piston was installed to determine whether piston surface conditions may effect combustion via fuel absorption into, and evaporation and burning of fuel out of the coating during the cycle. The hypothesis that this may be causing extended heat release durations in LHR engine has been presented by several experimenters and appears plausible. Unfortunately, this coating failed after a little more than one hour of test time with no relevant data collected. Upon closer inspection, 30% of the coating had delaminated from the center of the piston dome. It is the intent to investigate piston coating surface effects (porous vs. non-porous coatings) on combustion in LHR engines at some later date.

SUMMARY

For the particular LHR engine tested, whose configuration consisted of head cooling, no liner cooling and zirconia coated piston, test results are summarized as follows:

1) Combustion diagram analysis revealed that ignition delay was shorter, premixed fraction was less and heat release duration greater for the LHR engine. The volumetric efficiency was less for the LHR engine however, differences were not substantial (3% or less). Exhaust temperatures were greater in all LHR cases.

2) LHR engine performance results were mixed. Generally, at full load conditions the indicated specific fuel consumption was better (3%) at the high speed conditions and worse (5%) at the low speed conditions for the highest coolant, oil and intake manifold temperature test conditions. This corresponds directly with the higher instantaneous peak combustion pressures measured at the high speed conditions. Also as the head coolant temperatures were increased at the rated 2600 rpm condition, the trend indicated better fuel consumption. Part load data indicated equal specific fuel consumption at 2600 rpm and lower fuel economies (5% and 8%) at the low and mid speeds respectively.

3) TACOM finite element analysis predicted a 200°F higher peak temperature in general across the head firedeck for the LHR engine. The peak temperature was 813°F between the exhaust valves. Previous test work showed an increase of 250°F for the liner top ring reversal temperature and 200°F for the piston dome in an LHR engine.

CONCLUSIONS

1) Improved mixing of fuel and air is required in order to permit low heat rejection combustion to be optimized. More specifically, improved time phasing of premixed and diffusion burning must be accomplished if a desirable fundamental cycle heat release is to be derived. In order that the true overall advantages of low heat rejection diesel engines be met, continued combustion optimization of these systems must be pursued.

- 2) Heat release in LHR engines exhibit a far greater sensitivity to small changes than in normally cooled engines.
- 3) While experimentation verification is required hypotheses which identify the importance of seal coats on porous ceramic coatings to reduce cyclic thermal inertia appear to be plausible.
- 4) The thermal state of the test engine in-cylinder components, even at modest LHR conditions must be carefully analyzed at the higher temperature conditions in order to obtain acceptable performance, and structural and tribological integrity.
- 5) Continued use of fundamental combustion data such as generated within this paper, provide an indispensable ingredient for improving state-of-the-art burning laws within diesel cycle simulation programs.

References

- 1) Bryzik, W., and Kamo, R., "TACOM/Cummins Adiabatic Engine Program," SAE Paper 830314, SAE Trans. Vol 92, 1983.
- 2) Bryzik, W., and Kamo, R., "TACOM/Cummins Adiabatic Engine Program," SAE Paper 840428, Feb 1984.
- 3) Assanis, D., Wiese, K., Schwarz, E., and Bryzik, W., "The Effects of Ceramic Coatings on Diesel Engine Performance and Exhaust Emissions," SAE Paper 910460, Feb 1991.
- 4) Reddy, C. S., Domingo, N., and Graves, R. L., "Low Heat Rejection Engine Research Status: Where Do We Go from Here?" SAE Paper 900620, Feb 1990.
- 5) Borman, G. L., Farrell, P. V., Foster, D. E., Martin, J. K., and Myers, P. S., "Development of a Modular Transient Cycle Analysis Program for the Adiabatic Diesel and Other Compound Diesel Engines," University of Wisconsin, TACOM, RDT&E, Technical Report No 13558, pp. 57-71, June 1991.
- 6) Brands, M., Hoenhe, J., and Moore, C., "TACOM Adiabatic Engine Program Review," 4 Oct 1984.
- 7) Alkidas, A.C., "Performance and Emission Achievements with an Uncooled, Heavy Duty Single Cylinder Diesel Engine," SAE 890144, Feb 1989.
- 8) Gatowski, J. A., "Evaluation of a Selectively Cooled, Single Cylinder, 0.5 L Diesel Engine," SAE 900693, Feb 1990.
- 9) Dickey, D. W., "The Effect of Insulated Combustion Chamber Surfaces on Direct Injected Diesel Engine Performance, Emissions and Combustion," SAE 890292, Feb 1989.
- 10) Morel, T., Wahiduzzaman, S., and Fort, E. F., "Heat Transfer Experiments in an Insulated Diesel," SAE 880186, Feb 1988.
- 11) Woods, M., Schwarz, E., and Bryzik, W., "Heat Rejection from High Output Adiabatic Diesel Engine," SAE 920541, Feb 1992.

CARTLAND, VENEY, NICKOLAISEN, and WITTIG

Time-Resolved Infrared Diode Laser Spectroscopy:
A State Specific Probe of Atom-Molecule Reaction Dynamics

Harry E. Cartland*, CPT and David W. Veney, MAJ

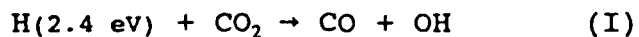
Photonics Research Center
United States Military Academy, West Point, NY 10996-5000

Scott L. Nickolaisen, Dr. and Curt Wittig, Prof.

ARO Center for the Study of Fast Transient Processes
University of Southern California, Los Angeles, CA 90089-0482

1. INTRODUCTION

Simple atom-molecule reactions such as are studied here



are of interest because of the role their surfaces play in combustion and atmospheric chemistry. In addition, their relative simplicity makes them amenable to computational studies, allowing a rigorous comparison between experiment and current dynamical theories.

An extensive literature already exists on reaction (I), and has recently been reviewed.¹ Most previous studies have focussed on laser-induced fluorescence (LIF) detection of the OH fragment, and include measurements of product state distributions and a determination of CM translation using sub-Doppler techniques. However in one instance, CO($v=0,1$) distributions have been measured under conditions comparable to ours using VUV LIF.² Also of note is the *ab initio* HOCO potential energy surface of Schatz et al., which was parametrized for the experimentally determined barrier of 9110 cm^{-1} .³

In contrast, reactions (II) and (III) have not been the subject of such thorough scrutiny. Earlier studies here have aimed at measurement of rate constants, yielding activation barriers for these reactions of 1350 cm^{-1} and 1540 cm^{-1} , respectively.^{4,5}

Through the application of time-resolved diode laser spectroscopy, our goal is to develop a detailed understanding of the factors influencing reactivity and energy disposal in this important class of abstraction reactions. These factors include collision energy, thermochemistry, and most importantly the potential energy surface connecting reactants to products.

2. EXPERIMENTAL SECTION

The UV laser pump/IR diode laser probe technique has been described elsewhere in detail.^{1,6} The apparatus used for the study of (I) was not identical to that used for (II) and (III), but the differences were for the most part minor. We provide a general description of the technique, adding specific details as necessary.

Translationally "hot" H atoms were generated in a 2 m pressure stabilized flow cell by UV laser photolysis of an appropriate precursor; H_2S (193 nm), HI (278 nm), and NO_2 (355 nm) for reactions (I), (II) and (III), respectively. The pulse energy, 10-150 mJ depending on the experiment, was monitored with a calibrated photodiode to allow normalization of the measured nascent product distributions. The premixed sample consisted of 20% of the appropriate atom source in neat reactant, and was maintained at 200-400 mTorr to minimize collisional relaxation of the nascent product.

A collinear CW IR diode laser then state selectively probed the $\text{CO}(v,J)$ reaction product. The output of the IR diode was collimated and passed through a monochromator to select a single laser mode with a nominal linewidth of $\approx 0.0003\text{ cm}^{-1}$. Because of anharmonicity, virtually any $\text{CO}(v,J)$ level is accessible by temperature and current tuning of the diode output, though more than one diode is usually required to cover the wavelength range of interest. The distributions reported here are based on a sampling of even and odd J states in both the P and R branches.

Prior to entering the sample cell, $\approx 10\%$ of the IR beam was split off and passed through a reference flow cell. The reference cell and a Ge etalon permitted precise wavelength calibration. Passing a discharge through the cell generated a reference population in higher vibrational states ($v \leq 9$),

allowing excited product states to be probed in the sample cell.

Standard line locking techniques were used to prevent drift of the diode laser. The reference detector signal was input to a lock-in amplifier. By using phase sensitive detection, a small dither in the diode output produced a first derivative error signal, which was fed back into the diode current controller to drive the laser output to line center.

The sample and reference beams passed through long wavepass filters ($\lambda > 2.5 \mu\text{m}$) before being focussed onto $\text{N}_2(l)$ cooled InSb detectors. The signal level was boosted by impedance matched preamplifiers and line drivers. The reference signal was input directly to the lock-in amplifier, while the sample signal underwent another stage of amplification and filtering prior to being input to a transient digitizer. The digitizer recorded the absorption transients at $10^8 \text{ samples s}^{-1}$, and hard wire averaged them in the CAMAC crate controller. Between 200 and 1500 transient signals were averaged at a repetition rate which was empirically adjusted to avoid sample depletion effects. The averaged data was stored, processed, and displayed on a PC AT computer using original software written in ASYST. Using a photodissociation process that produced CO, the detection system risetime was measured at $\approx 80 \text{ ns}$.

For time-resolved $\text{CO}(v,J)$ lineshape measurements, a fraction of the IR beam was directed to a vacuum spaced confocal etalon. Instead of dithering the diode laser wavelength, the etalon scanner driver was modulated thereby dithering the fringe positions in time. The line locking technique described above was used to slave the laser output to an etalon fringe, which was subsequently stepped across the absorption feature of interest. Simultaneous acquisition of a room temperature CO reference lineshape served to calibrate the step size in terms of wavelength.

3. RESULTS

Some typical $v=0$ signals from (II), the $\text{H}+\text{OCS}$ reaction, are shown in Fig. 1, where the raw data have been converted to absorbance units. These transient signals clearly illustrate the reaction and CO relaxation dynamics. The two component rise for low J is evidence of a small nascent population, increased by rotational relaxation, and to a lesser extent vibrational relaxation, as equilibrium is reached. In contrast, high J levels show substantial initial populations which rapidly decay.

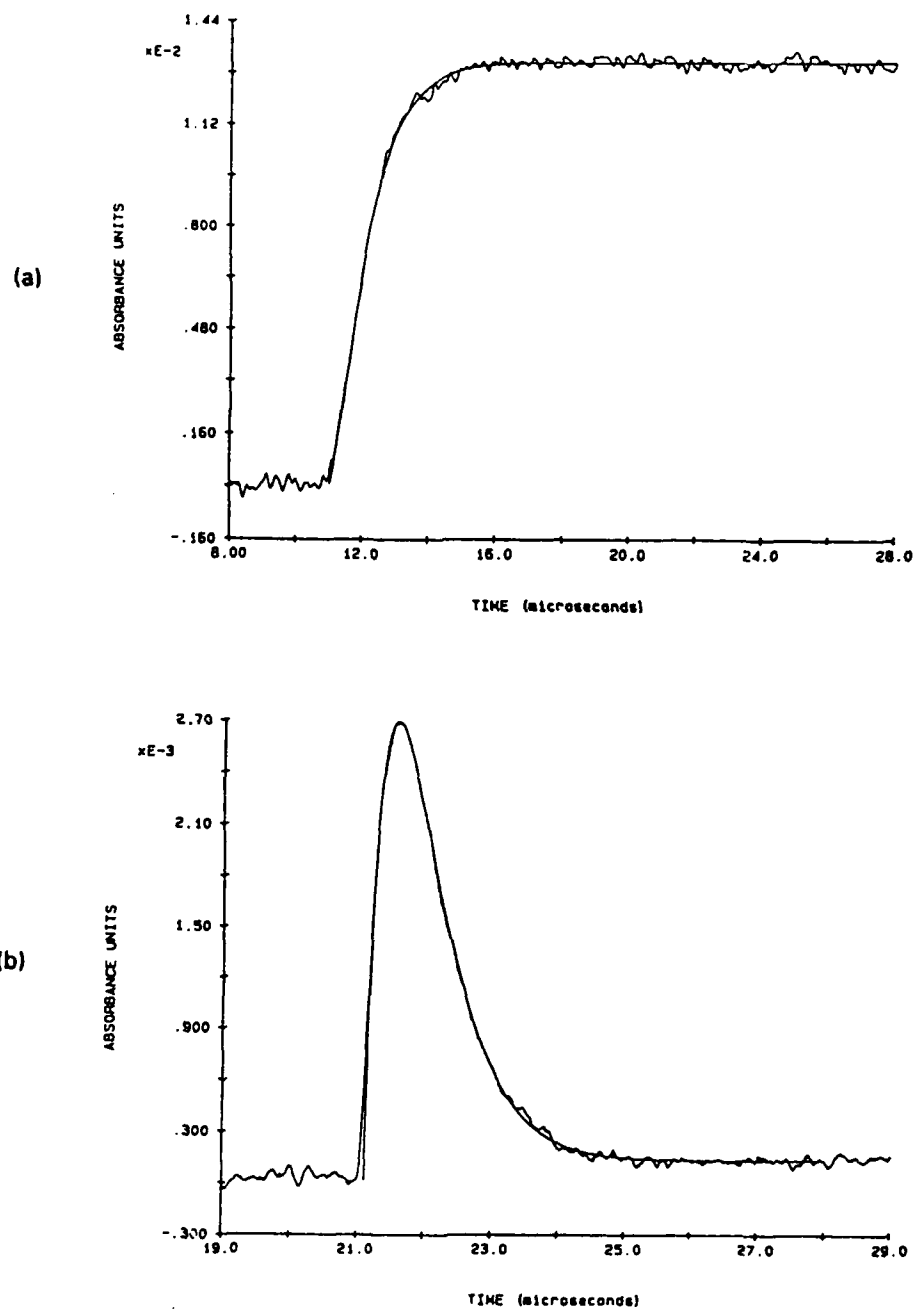
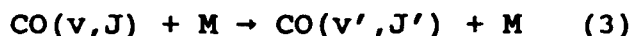
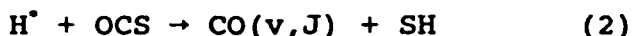


Figure 1 CO($v=0$) transient absorption signals from $\text{H}+\text{OCS} \rightarrow \text{CO}+\text{SH}$; (a) P(10), (b) P(35).

To interpret the data, we develop the following simple model, shown here illustrating (II), H+OCS, but of general applicability.



Reaction 1 is collisional deactivation, but effectively includes all processes that remove translationally "hot" atoms, such as diffusion out of the probed volume and various radical reactions. Reactions 2 and 2' represent reactive scattering into the monitored state $CO(v,J)$, and all other states $CO(v',J')$. Modelling the CO manifold as a quasi two level system, consisting of the probed state and a "bath" state with which energy is exchanged via reactions 3 and 3', yields an analytic expression for the time dependent $CO(v,J)$ population $[CO(v,J,t)]$. As the fits in Fig. 1 show, the transient absorption signals are well modelled by this simple scheme.

The derivation of the analytic expression for $[CO(v,J,t)]$ is discussed elsewhere in detail.^{1,6} Here we consider only two limiting cases. First, in the long time limit we find

$$[CO(v,J,t)]_{t \rightarrow \infty} = [H^*]_0 \cdot \frac{(k_2 \chi_{OCS} + k_2' \chi_{OCS})}{(k_1 + k_2 \chi_{OCS} + k_2' \chi_{OCS})} \frac{k_3'}{k_3 + k_3'}$$

where $[H^*]_0$ is the initial "hot" atom number density and χ_{OCS} is the mole fraction of the OCS collision partner. This expression has a straightforward interpretation: the concentration of $CO(v,J)$ after attainment of equilibrium is a product of the initial "hot" atom concentration, the branching ratio for reactive scattering, and the equilibrium fraction of molecules in that state. Of more interest, though, is the short time limit

$$[CO(v, J, t)]_{t=0} = k_2 \chi_{OCS} [H^*] \cdot [M] t$$

where $[M]$ is the total sample number density. Initially, the $CO(v, J)$ density is linear in time, with a proportionality constant dependent on the state specific rate constant k_2 . Therefore, a plot of the initial slopes of the transient signals gives a relative distribution of state specific rates, which is equivalent to plotting the actual nascent $CO(v, J)$ number densities.

Confidence in the distributions obtained in this manner requires that the model incorporate all essential features of the kinetics. In (II) and (III) for example, the unusually high relaxation rate of vibrationally excited CO in OCS ($P \approx 0.1$ for $CO(v=4)$) must be considered. Elsewhere, we explore the effect of processes specific to the individual systems.^{1,6} Inclusion of additional kinetics in the model of course improves the fit, but in no case will any of these secondary processes change the limiting expressions.

The time-resolved absorption technique used in our experiments is not as sensitive as MPI or LIF, but it does present one advantage. A straightforward distinction between gain and absorption clearly indicates the presence of population inversion, which raises an important point. Absorption techniques measure population differences. This means that the initial slope of the absorption signal is proportional to the difference of the upper and lower state specific rate constants. By probing successively higher vibrational levels until no further absorption is observed, the nascent rotational distribution of the highest populated vibrational level is directly determined. Working backwards, the difference measurements yield the distributions in lower vibrational levels. It is important to make these corrections, particularly for (II) where excitation of CO is observed up to $v=4$.

Using the initial slope method, CO rotational distributions were determined for all vibrational levels populated in (I), (II), and (III) within our detection limits. Some typical distributions are displayed in Figs. 2 and 3, and are discussed below.

Figure 2 shows a $CO(v=0)$ rotational Boltzmann plot for (III), $O+OCS$. Strictly speaking, Boltzmann distributions describe systems in equilibrium. This is certainly not the case for the nascent distributions reported here, but Boltzmann plots are often good approximations and can be quite

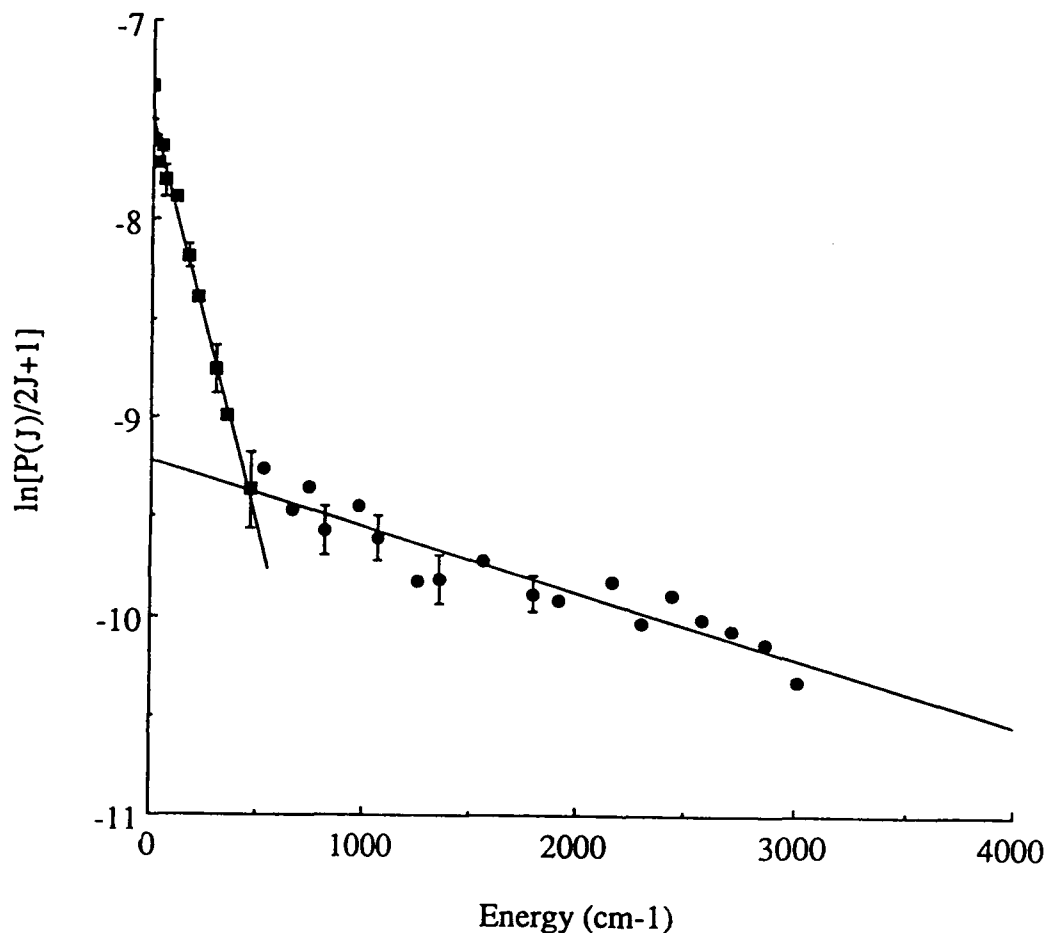


Figure 2 CO($v=0$) Boltzmann plot for O+OCS \rightarrow CO+SO. The low J portion is fit by $T_R = 350 \pm 20$ K, and the high J portion by $T_R = 4400 \pm 390$. Error bounds are $\pm\sigma$.

revealing, as the figure shows. The plot in Fig. 2 is clearly bimodal with a low J (energy) temperature $T_R=350$ K, and a high J temperature $T_R=4400$ K. Because the low J portion of the distribution has a temperature near ambient and the rise of the low J signals is delayed from the photolysis pulse, we attribute it to rotationally relaxed CO. Thus a rotational temperature $T_R=4400$ K is more representative of the nascent CO($v=0$) distribution. Relaxed CO absorption was also observed in $v=1$, yielding a limit on the vibrational branching ratio

$[v=1]/[v=0] \leq 0.05$, though poor S/N precluded determination of the nascent rotational distribution.

A better way to represent the data is a surprisal plot, or in terms of a surprisal parameter θ . The surprisal analysis procedure, developed by Levine, Bernstein, and coworkers,⁷ can be used to compare an observed rotational distribution $P(v,J)$ to a statistical maximum entropy distribution $P^*(v,J)$. Calculation of the "prior" distribution $P^*(v,J)$ involves summing over states and is well understood: the details are discussed elsewhere for systems of interest here.^{1,6} Surprisal plots are typically linear and, by definition, have positive slopes, or surprisal parameters θ_R , for distributions that are "colder" than statistical. A surprisal plot for CO($v=1$) from (I) is shown in the inset in Fig. 3, where the ordinate g_R is a reduced energy variable. The large positive rotational surprisal parameter $\theta_R=9$ indicates a substantial bias against rotational excitation in the CO product. The solid line overlying the rotational distribution in Fig. 3 represents the back-calculated distribution derived from the surprisal analysis, and fits the experimental data well. The CO($v=0$) rotational distribution was also found to be cold, $\theta_R=7$. Integration of the rotational distributions yields a vibrational branching ratio $[v=1]/[v=0] = 0.4$. For reaction (I), no signal was detected for 3 \leftarrow 2 transitions, indicating that little or no population of $v>1$ occurs.

Experimentally measured CO distributions for the three reactions studied are summarized in Table I.

Time-resolved sub-Doppler lineshape measurements on the CO($v=0$) product from (II) indicate substantial partitioning of energy into CM translation. A total of six transitions from $J=7$ to $J=30$ were examined: R(7), P(10), R(16), R(19), R(22), and R(30). A 3D plot for R(19) is shown in Fig. 4. Cross sections at 100 ns intervals were fit to Gaussian lineshapes, from which E_{CM} was determined. Since the S/N ratio is initially poor due to the small amount of product present, a preliminary estimation of the nascent E_{CM} was made by fitting a plot of E_{CM} versus time to a single exponential function, and extrapolating back to $t=0$. Such a plot for R(19) is shown in Fig. 4. In all cases the exponential fit appears to have underestimated the nascent CM kinetic energy. We take the average E_{CM} thus determined, 7180 cm⁻¹, as a lower limit for CO($v=0$).

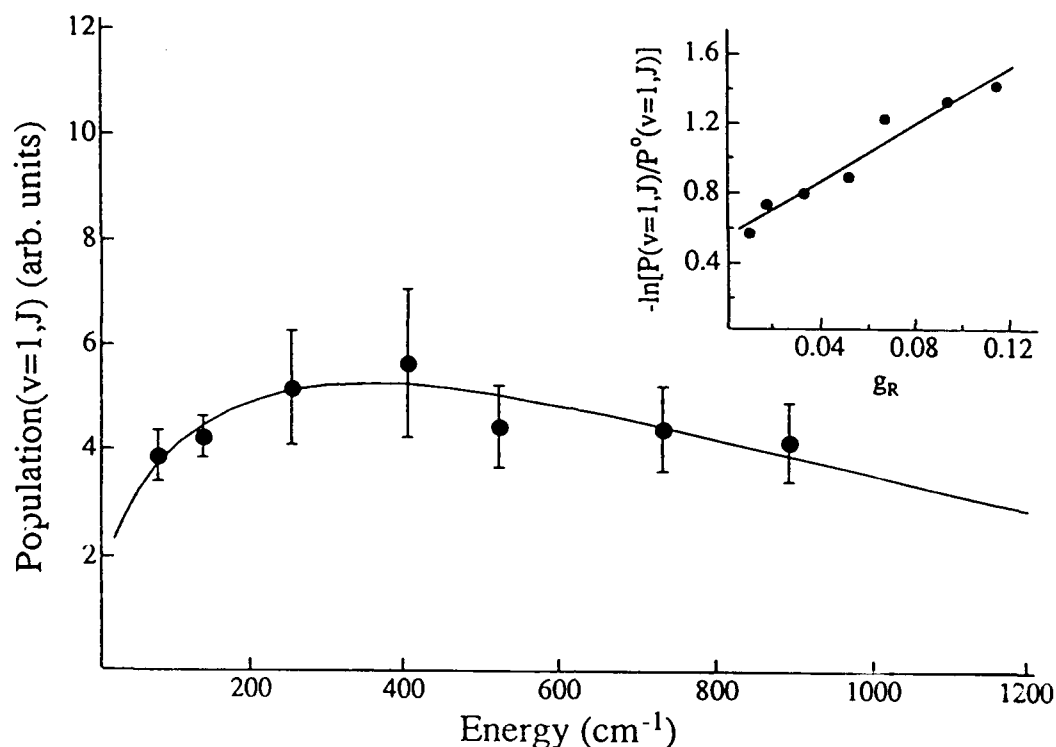


Figure 3 $\text{CO}(v=1)$ rotational distribution from $\text{H}+\text{CO}_2 \rightarrow \text{CO}+\text{OH}$. The solid line is the back-calculated distribution from the surprisal analysis (inset).

4. DISCUSSION

For reaction (I), $\text{H}+\text{CO}_2$, our results are consistent with those of Weston and coworkers,⁸ but disagree with the study of Rice et al. who found a $[v=1]/[v=0]$ ratio of 1.² Using VUV LIF, Rice et al. also reported an inverted $v=1$ rotational distribution peaked at $J=23$ which should have manifested itself as gain in our experiments: none was observed. While our observations are not in quantitative agreement with those of Rice et al., we both conclude that CO internal excitation is less than statistical. Previous studies of nascent OH distributions have shown that they too are colder than statistical,¹ while Chen et al. were able to demonstrate substantial partitioning into CM translation for $\text{OH}(v=0, \text{ low } J)$ using sub-Doppler techniques.⁹ In sum, these results

Table I Summary of nascent CO product distributions. Error limits are $\pm\sigma$.

Reaction	CO Vibrational Quantum Number	Relative Vibrational Population	Rotational Boltzmann Temperature T_R (K)	Rotational Surprisal Parameter Θ_R
H + CO ₂ → CO + OH	0	1.0	1500 ± 110	7.2 ± 0.7
	1	0.4	1030 ± 50	9.0 ± 0.6
H + OCS → CO + SH ^a	0	1.00	1400 ± 140	12.0 ± 1.4
	1	0.17	780 ± 45	21.3 ± 1.3
	2	0.10	590 ± 100	24.0 ± 3.8
	3	0.07	405 ± 20	28.6 ± 1.5
	4	0.03	175 ± 20	51.9 ± 5.6
O + OCS → CO + SO	0	1.00	4400 ± 390	3.7 ± 0.5
	1	≤0.05	-	-

^a $T_v = 3890 \pm 830$ K; $\Theta_v = 3.6 \pm 1.1$

indicate a propensity for CM translation at the expense of product internal excitation, though it is also important to note that both the picosecond spectroscopy studies of Zewail and coworkers¹⁰ and the surface calculations of Schatz et al.³ implicate a HOCO intermediate in this reaction.

Our picture of (II), H+OCS, is not so complete. Predissociation of SH precludes studying this fragment with LIF, however SD distributions in the analogous D atom reaction have been measured and are, like OH, found to be cold and nonstatistical.¹¹ By conservation of energy, our sub-Doppler lineshape measurements place an upper limit of 5900 cm⁻¹ for the amount of available energy channeled into the SH fragment. Again, this system displays an energy disposal propensity towards CM translation. Unfortunately there is no HOCS potential surface to aid in interpretation of the experiments, but the data are consistent with a near collinear collision geometry for the reactive process.⁶

Least well understood is reaction (III), O+OCS. Further experimentation is necessary to elucidate the dynamics of this reaction. SO product distributions must be measured, and a determination of the extent to which energy is partitioned

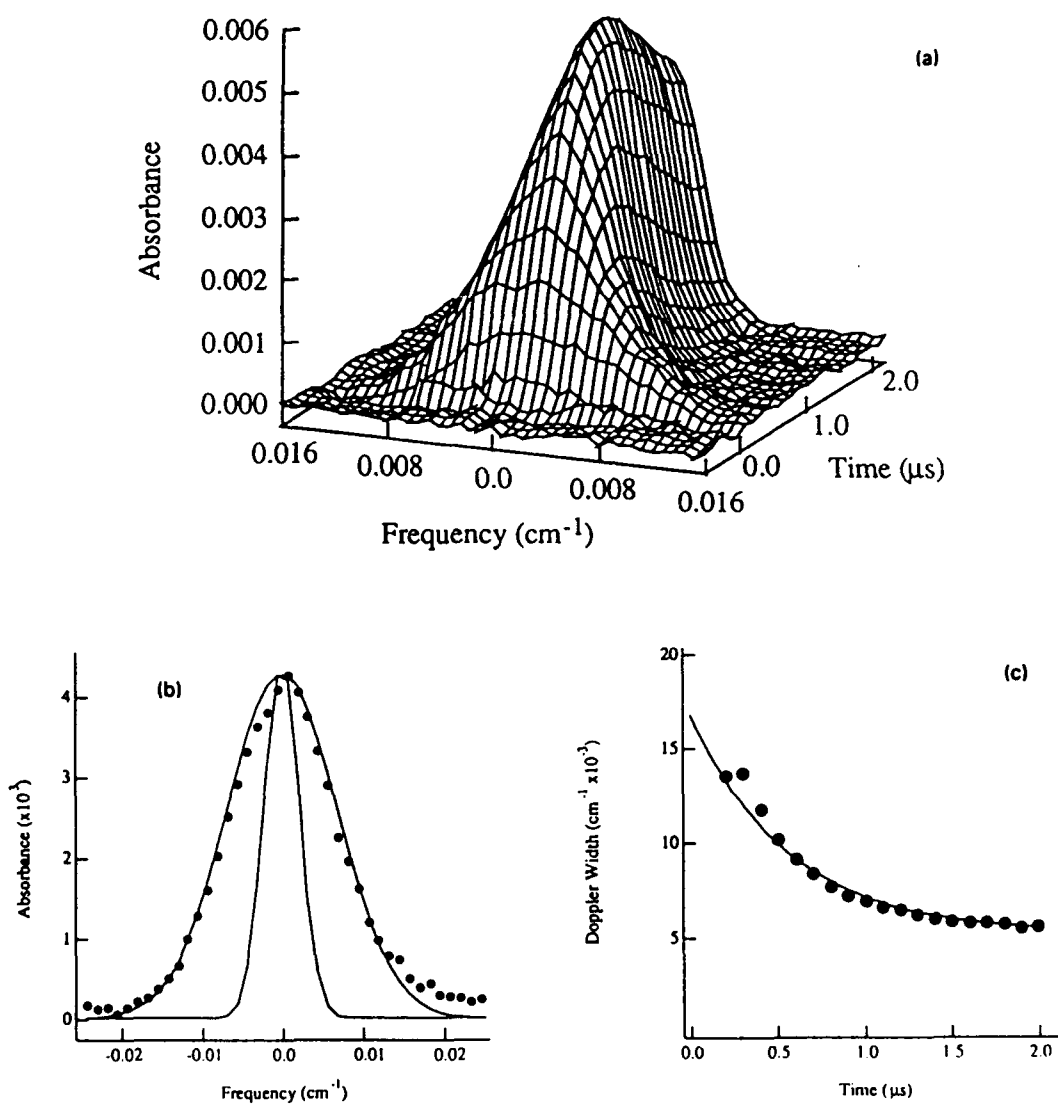


Figure 4 Sub-Doppler spectroscopy of $\text{CO}(v=0, R(19))$ from $\text{H}+\text{OCS} \rightarrow \text{CO}+\text{SH}$; (a) absorbance as a function of frequency and time, (b) lineshape, Gaussian fit, and 298 K reference lineshape at 500 ns, (c) Doppler width.

into CM translation must be made. A OSCO potential surface would be useful in addressing the role of reaction intermediates. Despite the relative dearth of information on (III), the data obtained thus far are consistent with the trends set by (I) and (II); CM frame translation at the expense of product internal excitation.

Factors pertinent to a discussion of the dynamics for these systems are listed in Table II. With a total of $\approx 10,000$ to $20,000 \text{ cm}^{-1}$ available to the products, the fraction channeled into CO internal degrees of freedom is surprisingly small, $\approx 15\%$. Although the data set comprises a wide range of collision energy, activation barrier, and reaction enthalpy, the nascent CO vibrational and rotational distributions are found to be cold when compared to statistical theory. Thus intramolecular vibrational redistribution (IVR) is not complete along the reaction coordinate, nor, as the data in Table I indicate, has equilibration occurred between rotational and vibrational degrees of freedom.

The data suggest the following. First, only a fraction of the total energy "pool" of the system is available for internal excitation of the product fragments, or the phase space for the reactive process is somehow restricted. Second, in that fraction available for product excitations, energy is conserved to the extent that partitioning into vibrational degrees of freedom occurs at the expense of product rotation, or the absence of coupling between rotations and vibrations causes rotational temperature to drop with increasing vibrational quantum number. And third, the majority of the available energy is repulsively released in the exit channel, and is directed primarily into relative translational motion of the separating fragments. Further study of the potential energy surfaces of these atom-molecule systems may reveal the dynamical source of the observed nonstatistical behavior.

For reactions (I) and (II), because of the large difference in mass, it is unlikely that the H atom will simply abstract an O or S atom in passing. While the required slowing of the H atom to form a chemical bond implies the existence of an intermediate, it does not guarantee statistical distribution of available energy among product degrees of freedom. In the energy regime of these experiments, $\approx 10^4 \text{ cm}^{-1}$ at the transition state, RRKM estimates of reaction rates are on the order of 10^{13} s^{-1} .¹² Since IVR may not be complete on this time scale, the applicability of statistical theory to these experiments is questionable even though a local minimum exists on the surface.

Table II Potential surface features and partitioning of available energy.

Reaction	H + CO ₂ → CO + OH	H + OCS → CO + SH	O + OCS → CO + SO
CM Collision Energy E _{CM} (cm ⁻¹)	19170	11070	1650
Reaction Enthalpy ΔH ₂₉₈ (cm ⁻¹)	+8960	-3930	-17670
Activation Barrier E* (cm ⁻¹)	9110	1350	1540
Transition State Energy E _{CM} -E* (cm ⁻¹)	10060	9720	110
Available Energy E _{CM} -ΔH ₂₉₈ (cm ⁻¹)	10210	15000	19320
CO Internal Excitation E _{int,CO} (%)	15	13	16
XY Internal Excitation ^a E _{int,XY} (%)	23	<39	-
Translational Energy E _{CM} (%)	62	>48	-

^a XY = OH, SH, SO as appropriate

An explanation of nonstatistical behavior in reaction (III) presents a challenge of another sort. A lack of data makes interpretation difficult, but a different mechanism is almost certainly operative since the amount of energy over the activation barrier, 110 cm⁻¹, falls within the realm where statistical theory could apply. The reaction mechanism for (III) must explain a CO product rotational temperature that is significantly higher than its vibrational temperature, unlike (I) and (II) where the opposite is true. Two mechanisms are consistent with the observed partitioning. In the first, a short-lived OSCO intermediate forms with energy preferentially

deposited in a degree of freedom that becomes product rotation (incomplete IVR). In the second, a long-lived OSCO intermediate forms (IVR complete), and rotational excitation is the result of a torque on the CO fragment in the exit channel. It will be interesting to see if this O atom reaction shows the same propensity towards product fragment CM translation that the H atom reactions do: there is good indication that it will.

5. CONCLUSION

Using a sensitive and selective IR diode laser probe technique, we have studied reactive scattering of H and O atoms with CO, and OCS through state resolved monitoring of the CO product. The CO rotational and vibrational distributions obtained under arrested relaxation conditions are significantly colder than those predicted by statistical theory. Our results for these atom-molecule reactions, coupled with earlier studies, indicate incomplete IVR even if intermediates are present, with preferential partitioning of energy into CM frame translational degrees of freedom.

6. ACKNOWLEDGEMENT

This work was supported by the Army Research Office and the Office of Naval Research.

7. REFERENCES

1. S.L. Nickolaissen, H.E. Cartland, and C. Wittig, *J. Chem. Phys.*, accepted for publication.
2. (a) J.K. Rice, Y.C. Chung, and A.P. Baronavski, *Chem. Phys. Lett.* **167**, 151 (1990). (b) J.K. Rice and A.P. Baronavski, *J. Chem. Phys.* **94**, 1006 (1991).
3. G.C. Schatz, M.S. Fitzcharles, and L.B. Harding, *Faraday Disc. Chem. Soc.* **84**, 359 (1987).
4. (a) S. Tsunashina, T. Yokota, I. Safarik, H.E. Gunning, and O.P. Strausz, *J. Phys. Chem.* **79**, 775 (1975). (b) J.H. Lee, L.J. Stief, and R.B. Timmons, *J. Chem. Phys.* **67**, 1705 (1977).

5. (a) R. Atkinson, D.L. Baulch, R.A. Cox, R.F. Hampson Jr., J.A. Kerr, and J. Troe, *J. Phys. Chem. Ref. Data*, p. 881, Vol 18, 1989. (b) W.B. DeMore, D.M. Golden, R.F. Hampson, C.J. Howard, M.J Kurylo, M.J. Molina, A.R. Ravishankara, and S.P. Sander, *Chemical Kinetic and Photochemical Data for Use in Stratospheric Modeling, Evaluation Number 8*, JPL Publication 87-41, 1987.
6. S.L. Nickolaissen, *State Resolved Product Dynamics in Atom-Molecule Interactions*, PhD dissertation, University of Southern California (1991).
7. R.D. Levine and R.B. Bernstein, *Acc. Chem. Res.* 7, 393 (1974).
8. R.E. Weston, private communication.
9. Y. Chen, G. Hoffman, D. Oh, and C. Wittig, *Chem. Phys. Lett.* 159, 426 (1989).
10. N.F. Scherer, L.R. Khundkar, R.B. Bernstein, and A.H. Zewail, *J. Chem. Phys.* 92, 5239 (1990).
11. D. Hausler, J. Rice, and C. Wittig, *J. Chem. Phys.* 91, 5413 (1987).
12. G. Hoffman, D. Oh, Y. Chen, and C. Wittig, *Israel J. Chem.* (1989).

Dynamic Impact Response of Titanium Aluminide Composites

Mr. Ernest S. C. Chin*, Materials Producibility Branch (SLCMT-MEM)
and

Mr. Patrick Woolsey, Materials Dynamics Branch (SLCMT-MRD),
U.S. Army Materials Technology Laboratory, Watertown, MA 02172-0001

INTRODUCTION

Within the past 20 years, a wide range of composite materials have emerged to meet expanding requirements for high performance materials. The principle of consolidating two or more different materials to obtain a composite having properties unattainable with the respective monolithics is by no means new. One of the earliest instances of a composite material is cited in the Bible, where clay and straw are the critical ingredients in forming brick. In this case, the straw serves as a reinforcement which provides strength and rigidity to the clay matrix. Today, fibers, whiskers, and particulates can serve as abrasion-resisting agents, stiffeners, tougheners, or strengtheners in polymer, ceramic, metallic, and more recently, intermetallic matrices. Tailoring material properties through a composite concept may be the primary means of meeting future materials requirements. Researchers today have only scratched the surface of composites' full potential.

With the exception of fiber reinforced polymers, only a limited amount of data is available concerning the ballistic behavior of composite materials. Polymer matrix composites are used in a number of ballistic protection applications. Fiber reinforced polymer spall liners are utilized to capture perforation debris from impacts, and in structural components with a secondary ballistic defeat function. Limited studies on aluminum matrix composites have indicated that reinforced aluminum was better¹ or significantly more efficient² than the monolithic base alloy against tungsten penetrators. Preliminary studies at the U.S. Army Materials Technology Laboratory (AMTL) have demonstrated significant improvement in penetration resistance against armor piercing (AP) projectiles through thermomechanical processing of a silicon carbide reinforced 6061 aluminum system³. Further studies on metal matrix composites are being pursued at AMTL and other research facilities. However, there are no known ballistic studies on the performance of intermetallic materials or intermetallic composites available in the open literature today.

Traditionally, intermetallic compounds have been perceived mainly as precipitation hardening and strengthening agents in metallic alloys. Due to the recent thrusts by the DoD (Department of Defense)/ industrial IHPTET (Integrated High Performance Turbine Engine Technology Initiative), NASA HITEMP (High Temperature Engine Materials Technology Program), and the development plans for the NASP (National Aerospace Plane), the engineering community has acknowledged that traditional monolithic metals can no longer meet all the high temperature and low density criteria required. With the potential of intermetallics as high temperature engineering materials recognized, intense research and development efforts aimed at exploiting their capabilities began. Up to the present, titanium aluminides have been the most studied intermetallic system, and they are still the focus of numerous research efforts^{4,5}.

Low density, high temperature stiffness retention, creep and oxidation resistance are some of the valuable properties exhibited by titanium aluminides. The critical obstacles to the broad utilization of titanium aluminides in engineering applications are poor room temperature (RT) ductility and toughness associated to this class of material. Significant progress in addressing and improving ambient temperature ductility and fracture toughness is being made through suitable alloying, thermomechanical processing (TMP), and composite reinforcement routes.

Advanced ceramic materials offer the armor system designer a high efficiency option to traditional metallic armor materials. Ceramics commonly considered for armor applications share a number of features qualitatively similar to those exhibited by titanium aluminides. Low ductility, low toughness, low density, excellent high temperature property retention, good corrosion and oxidation resistance⁶ are some common characteristics found in intermetallics and ceramics. However, there are a number of drawbacks associated with studying ceramics for dynamic impact applications. Ballistic impact events occur at strain rates on the order of 10^4 to 10^5 (for projectiles at ordnance velocities). Magnifying the already difficult task of designing with brittle ceramic materials is the problem that the conditions of dependence on dynamic yield/failure criteria are not yet satisfactorily understood. Also complicating this situation, postmortem examination of failed ceramic material is hindered by the severe comminution of the impact region. Given the above property similarities, together with the readily performed means for modifying titanium aluminides as opposed to the difficulty of producing such changes in ceramics, it is of interest to see what behavior these intermetallic materials may exhibit under conditions of ballistic impact.

Titanium Aluminide Composites

Two categories of intermetallic matrix composites (IMC) are being developed today: fiber reinforced IMC and particulate reinforced IMC. Fiber reinforced intermetallic composites provide excellent high temperature stiffness and strength with good damage tolerance. The reinforcements typically require a tailored surface chemistry to provide high temperature stability and interfacial compatibility. Consolidation of the product can be performed by employing powder metallurgy processes, or through a spray deposition and hot press route. Either of these fabrication methods is relatively expensive when compared to conventional ingot metallurgy. Recently, tremendous advances have taken place in the development of *in situ* composites. Rather than relying on mechanical blending to incorporate reinforcements, selected dispersoids are precipitated by an exothermic reaction in the melt through a conventional ingot metallurgy route. Martin Marietta Laboratories' (MML) proprietary XD™ metallic⁷ and intermetallic⁸ composites are produced by a process based on the aforementioned principle. As a result of this advantage, it was decided to employ materials produced with this technology for the ballistic study.

The XD™ intermetallic composites examined in this study are TiB₂ reinforced titanium aluminides with a composition near the gamma(γ) region of the Al-Ti phase diagram⁹ (see Figure 1). In the XD™ titanium aluminide composites, the dispersoids are typically single crystal¹⁰, and act as nucleation sites in refining and stabilizing the as-cast microstructure¹¹. A fine grained microstructure provides good room temperature ductility. At elevated temperatures, the reinforcements pin dislocations to retard creep flow, thus allowing excellent mechanical property retention¹². The reinforcement loading, distribution, and size play a major role in affecting all mechanical properties exhibited by this material. For instance, a reinforcement loading of up to 7 vol% can increase the yield strength from the 400 MPa characteristic of the matrix to around 600 MPa with no significant loss of ductility. The ductility in XD™ titanium aluminides for this level of reinforcement ranges from 0.8 to 1.1%. Both direct and

indirect strengthening mechanisms contribute to this effect (private communication from Dr. Steven Kampe, Martin Marietta Laboratories, 1992). The particulars of the titanium diboride reinforcement distributions in the titanium aluminide composite can be tailored and established during the XD™ casting process.

The matrix selected for this study has the nominal composition Ti-48 at.% Al. This composition corresponds to the portion of the phase diagram with two reactions: a peritectic ($L + \alpha \rightarrow \gamma$) and an eutectoid ($\alpha \rightarrow \alpha_2 + \gamma$). Three prominent microstructures have been characterized in near gamma titanium aluminide monolithics: the single phase gamma structure; the $\alpha_2 + \gamma$ lamellar structure; and the duplex structure¹³. Variations and combinations of these microstructures are achievable through heat treatment and thermomechanical processing. The lamellar structure consists of grains having alternating layers of ordered face-centered tetragonal (L1₀) γ (TiAl) and ordered hexagonal (DO₁₉) α_2 (Ti₃Al) phases (see Figure 2a). A non-equilibrium lamellar structure is obtained by cooling directly from the liquid region. A fully lamellar structure is typically achieved by cooling after heat treatment in the high temperature alpha single-phase field. The α_2 and γ lathes have strong crystallographic orientation relationships toward each other ($(0002)\alpha_2 \parallel (111)\gamma$; $[1120]\alpha_2 \parallel [110]\gamma$)¹⁴. Heat treating a near- γ alloy in the $\alpha + \gamma$ temperature range will promote the transformation of the lamellar structure toward a more equiaxed structure. Gamma grains tend to nucleate and grow in the lamellae and at grain boundaries. The resulting microstructure is typically referred to as the subtransus microstructure (see Figure 2b). The duplex structure consists of lamellar grains and recrystallized grains of α_2 and γ . This microstructure is commonly found in forgings and extrusions heat treated in the $\alpha + \gamma$ phase field. At room temperature, the lamellar structure is preferable for toughness where as the duplex structure is more preferable for ductility¹⁵. With a subtransus microstructure, ductility increases over that of a lamellar structure as the nucleating γ grains multiply and grow. Corresponding results were found in XD™ titanium aluminide composites¹⁶. The focus of this study is to perform a limited study on the effects of microstructure on dynamic impact. The wide range of microstructure and properties available with the XD™ titanium aluminide composites provides an excellent opportunity to study this effect.

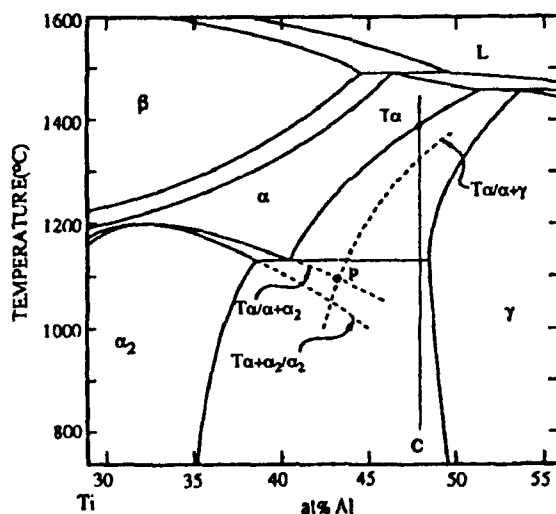


Figure 1. Al-Ti Phase Diagram [from ref. 9]

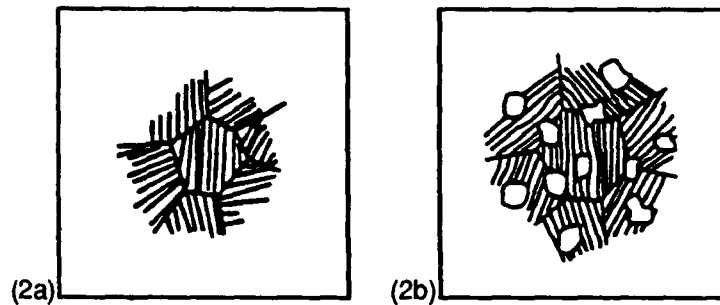


Figure 2. Illustration of a Lamellar Structure (a), and a Subtransus Structure (b) [from ref. 9]

Residual Penetration Test Method

In order to conduct the ballistic evaluation, it was decided to employ a method for which a large database of ceramic performance results is readily available, and which would produce results with a minimum number of tests to make best use of the limited material. The method selected was residual penetration ballistic testing. This is a relatively new technique for the evaluation of armor ceramic materials, which has gained considerable acceptance as a performance screening test within the armor materials community over the past several years. An early use of the methodology as a measure of performance for determining efficiency of applique armors was described by Yaziv, Rosenberg, and Partom¹⁷. Residual penetration tests were also employed by Rosenberg, Bless *et al.*¹⁸, and Rosenberg and Yeshurun¹⁹ to evaluate ceramics against small arms projectiles. An independent effort at the U.S. Army Materials Technology Laboratory to develop a ballistic test technique for performance ranking of ceramics against long rod penetrators resulted in a residual penetration test which has been described by Woolsey *et al.*^{20, 21, 22}. Other recent work using similar configurations has been described in papers by Alme and Bless²³, Bless *et al.*²⁴, and Anderson and Morris²⁵.

Residual penetration tests rely on the construction of a performance map for a material over a range of areal densities. Areal density is the measure of target material mass per unit surface area. Varying the areal density corresponds to varying the thickness of target material. Performance is measured by the depth of penetration of a projectile into a semi-infinite backplate after passing through a ceramic target. The penetration in the backplate after the target is considered to be the residual penetration. For convenience, this value is often referred to as "DOP", for "depth of (residual) penetration," in the AMTL implementation of residual penetration testing. This method, which has been adopted as the standard kinetic energy penetrator screening test for ceramic materials by both the U.S. Army Ballistic Research Laboratory and the AMTL, was the one employed to perform all ballistic tests for this paper. The basic assumptions which must be satisfied to ensure applicability of the test are as follows: that the projectile class of interest is a ductile heavy alloy long rod; that the operative penetration mechanism is hydrodynamic erosion, rather than rigid body penetration, and that the backing plate is of effectively infinite thickness with regard to the final position of the penetrator, so that no rear surface effects influence the penetration. Material comparisons may be made between selected baseline materials, for which a performance map has been developed, and candidate materials tested at any single areal density. This test method offers a number of advantages over traditional ballistic limit test methods by eliminating factors introduced by non-standard test configurations, targets with low ceramic percentages, and backplate effects. It also readily allows parametric analysis of the effects on material performance resulting from differing material properties.

EXPERIMENTAL PROCEDURES

Standard Ballistic Test Conditions for Ceramics

The standard residual penetration target consists of a material confinement package in contact with a thick steel backplate. The standard target package used is illustrated in Figure 3a and 3b. It consists of a 15 cm by 15 cm square ceramic tile of a selected thickness, which is laterally confined in a steel frame with a 2.5 cm web and a depth equal to or greater than the tile thickness. Ceramic tile thicknesses used in these tests have ranged from 0.625 cm to 10 cm. No cover plate is employed. An epoxy resin is used to retain the tile in the frame. In addition, a thin layer (approximately 0.5 mm) of epoxy is spread behind the tile to accommodate any slight surface irregularities, thus preventing nonuniform contact between the ceramic and the backing. A 0.5 mm aluminum sheet is used behind the tile to contain the epoxy within the frame. Backup blocks are monolithic rolled homogeneous armor steel (commonly termed RHA), per U.S. MIL-A-12560, Class 3, in 10 or 12.5 cm thickness. Specification requirements for this steel mandate the same hardness range for these thicknesses. Ceramic packages are mechanically clamped to the backup block for testing (as shown in Figure 3c.).

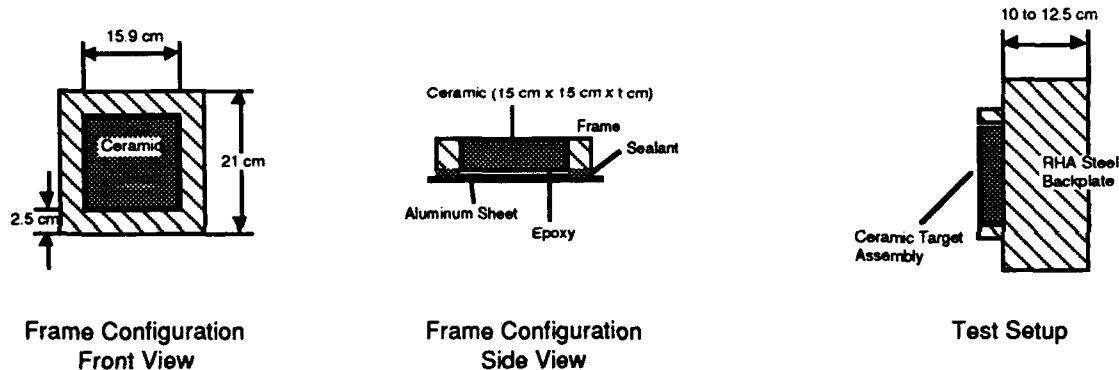


Figure 3 a-c. Ceramic Target Assembly

The standard projectile used is a 91% tungsten-nickel-iron long rod penetrator with an aspect ratio (L/D) of 10:1 and an hemispherical nose. Rod weight is 65 grams. The tungsten alloy employed is a standard commercial material, with a density of 17.45 g/cc. Penetrators are launched in a discarding sabot package from a powder gun fitted with a 20mm smoothbore test barrel. The nominal or standard velocity desired for all ceramic shots is 1494 m/s (4900 ft/s). This velocity was originally chosen in order to produce the maximum practical residual penetration while being consistently achievable under operating conditions. Measurement of projectile yaw and velocity are determined from flash X-ray images taken of the rod in flight, using a 150 kV Hewlett-Packard X-ray system, which drives two pairs of orthogonal tube heads.

All residual penetration measurements are taken directly from sections of the RHA steel block. Penetration cavities are directly sectioned, and measurements are made on the sections as indicated in Figure 4. Residual penetration is determined as $DOP = T_b - a$. This approach avoids inaccuracies which would arise from a change in local block dimension due to the plastic yielding of the surface immediately surrounding the penetration cavity.

At the chosen test velocity, the reference penetration P_0 is 68.6 mm for the rod employed. A more complete discussion of the penetration performance of the tungsten alloy rods used in the AMTL test has been published by Dowding et al.²⁶ Papers by Hohler and Stilp²⁷, and Anderson and Walker²⁸, among others, provide detailed analyses regarding the penetration of long rods in homogeneous semi-infinite media.

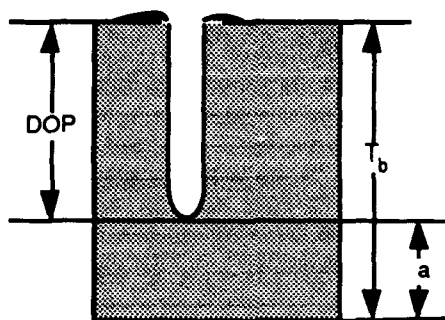


Figure 4. Residual Penetration Measurement

A variety of different ceramics have been tested at AMTL in order to provide a suitable basis for material comparisons, and to determine the validity of the test procedure. A number of reference ceramics of varying classes, for which complete residual penetration maps have been developed, are used as the bases for comparisons to materials being tested. In general, one attempts to select a reference material having a bulk density close to that of the candidate material. The reference material selected for this paper is a 90% pure, sintered aluminum oxide from Coors Ceramics (AD90). A comprehensive performance map is available for this material, consisting of over 50 tests across a wide range of areal densities. A linear fit to the data produces a simple formula whereby the DOP may be determined at any given areal density as follows:

$$\text{DOP (mm)} = 63.977 - (0.19071 \times \text{Areal Density (kg/m}^2\text{)})$$

The variation in DOP for the AD90 reference data at a constant areal density is 1.3 mm within one standard deviation. This is comparable to the variation observed in the steel target alone.

Material Conditions

The titanium aluminide IMC's used for this study were produced by MML via the XD™ process, and provided through the Naval Research Laboratory (NRL). Materials studied included 5, 10, and 15 vol% TiB₂ reinforced Ti-48 at.% Al castings. In addition, a 7.5 vol% TiB₂ reinforced Ti-48Al-2V forged pancake was provided by MML. For simplicity, the 5, 7.5, 10, and 15 vol% TiB₂ reinforced Ti-48 at.% Al will be referred to in the text as 5XD, 7.5XD, 10XD, and 15XD respectively. Chemical analyses of 5XD, 10XD and 15XD are shown in Table I (data furnished by Dr. C. Robert Crowe, Naval Research Laboratory, 1992). The 7.5XD was alloyed with 2 at.% V, whereas the 5XD, 10XD, and 15XD were all nominally binary alloys.

The 5XD, 10XD, and 15XD were 15.5 cm diameter ingots in the as-cast condition whereas the 7.5XD was a 3.81 cm thick hot forged pancake. The as-cast ingots had a lamellar structure developed during the solidification process. The relative grain size of the various ingots decreased as the reinforcement loading increased (see Figures 5a-c). The microstructure of the forging was a fine

duplex structure with lamellar and equiaxed recrystallized $\alpha + \gamma$ grains. Due to the thermomechanical processing, the forging had the finest grain structure among the composites (see Figures 5d-e).

Table I. Chemical Composition of TiAl Composites

Normal Composition	Ti	Al	TiB ₂	O ₂	N ₂	H ₂	C
5XD (Ti-48 at.% Al + 5 vol% TiB ₂)	51.4	48.6	5.1	0.096	0.007	0.0016	0.008
10XD (Ti-48 at.% Al + 10 vol% TiB ₂)	51.5	48.4	10.4	0.157	0.011	0.0017	0.008
15XD (Ti-48 at.% Al + 15 vol% TiB ₂)	52.6	47.4	14.4	0.178	0.016	0.0013	0.019

Notes: Ti and Al concentrations in atomic percent (at.%)

TiB₂ concentration in volume percent (vol%)

O₂, N₂, H₂, and C concentrations in weight percent (wt.%)

Table II. Vickers Hardness (1 kg Load)

	As Received	Heat Treated at 1330°C for 20 Hours
5XD (Ti-48 at.% Al + 5 vol% TiB ₂)	467.2	411.2
10XD (Ti-48 at.% Al + 10 vol% TiB ₂)	562.6	444.1
15XD (Ti-48 at.% Al + 15 vol% TiB ₂)	550.4	-
7.5XD (Ti-48 at.% Al-2V + 7.5 vol% TiB ₂)	414.7	505.1

In order to study the effects of ductility, selected specimens were heat treated toward a more ductile duplex structure. Half of the as-cast 5XD and 10XD, together with the forging, were annealed in the $\alpha + \gamma$ temperature region. Due to limited availability of material, the 15 XD was excluded from the heat treatment. The $\alpha + \gamma$ region of the Al-Ti phase diagram is bounded by the $\alpha \rightarrow \alpha_2 + \gamma$ eutectoid isotherm and the α transus (see Figure 1). Previous study of 5 vol% TiB₂/Ti-48 at.% Al determined the $\alpha \rightarrow \alpha_2 + \gamma$ eutectoid isotherm to fall between 1300°C and 1310°C, and the α transus temperature to fall between 1360°C and 1375°C²⁹. The heat treatment condition selected for both the as-cast ingots and the forging was 1330°C for 20 hours. These heat treatments were performed in a vacuum furnace, and all specimens were wrapped in tantalum foil. Hardness changes resulting from this heat treatment are found in Table II. A subtransus microstructure corresponding to a lower hardness was developed in the 5XD and 10XD ingots as a result of this heat treatment (see Figure 6). Coarsening of the duplex structure toward a subtransus microstructure was observed in the 7.5 XD after the heat treatment (see Figure 7). The resulting microstructure corresponded to an increase in hardness. It is generally accepted that as hardness increases, strength increases.

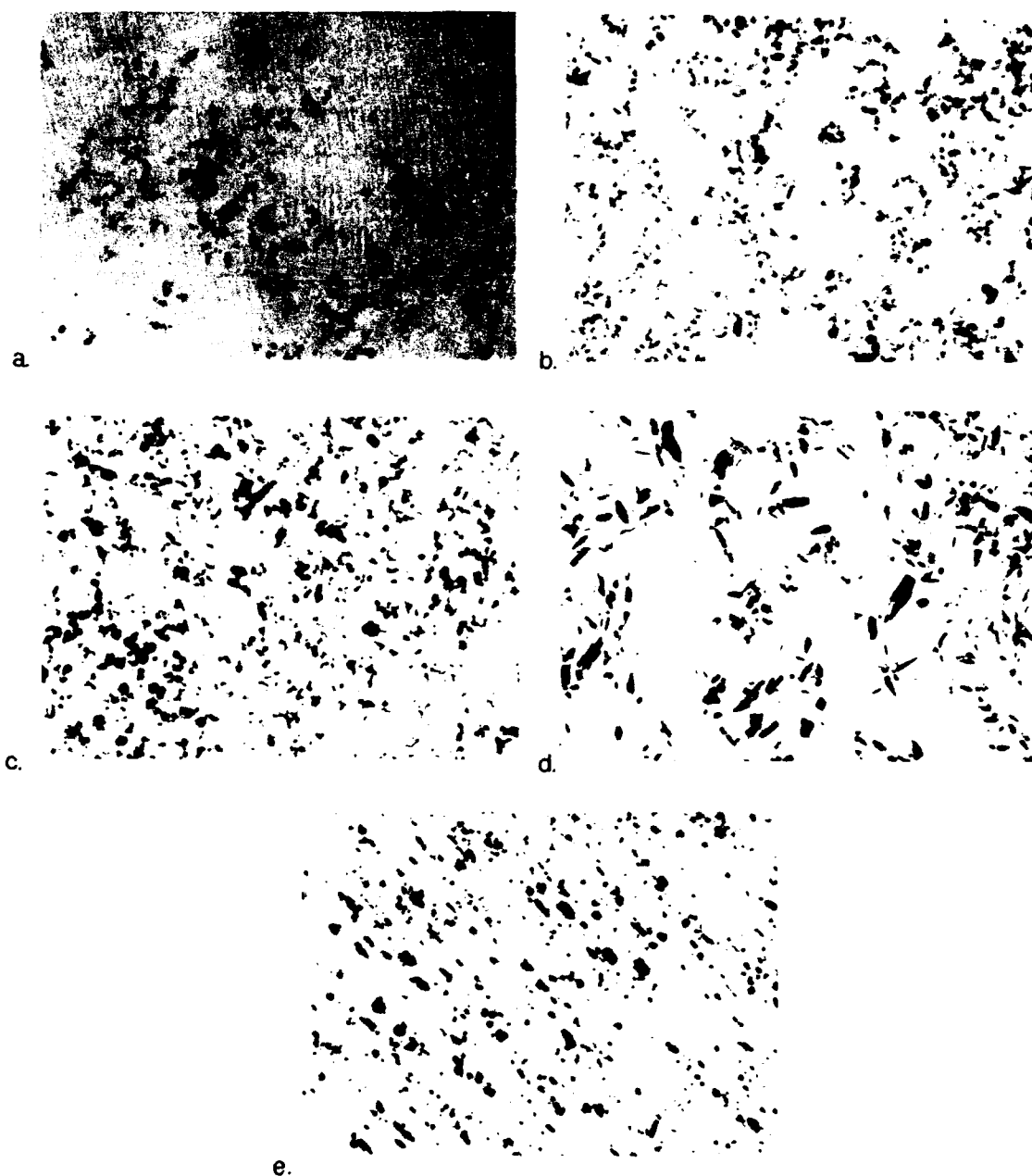


Figure 5. Backscatter Electron Images of As-Received 5XD(a), 10XD(b), 15XD(c),
Cross Section of 7.5XD(d), and Top Section of 7.5XD(e)
(Mag 200X)



Fig. 6. Post-Heat Treatment 5XD (Mag 200X)

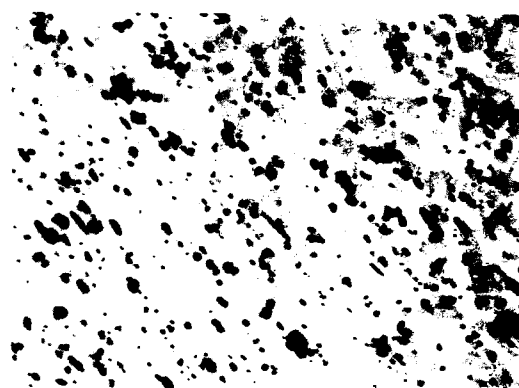


Fig. 7. Post-Heat Treatment 7.5XD (Mag 200X)

Ballistic Test Specimens

The ballistic performance of the XD™ titanium aluminide composites was evaluated by means of the residual penetration (DOP) test previously described. The as-received and heat treated ingots along with the heat treated forging, were sectioned to 2.54 cm and 5.08 cm thick plates by mean of electrode discharge machining (EDM). Due to the lateral size of the available ingots (15.5 cm diameter), it was not feasible to produce a square plate of the size used in the ceramic configuration. Therefore, confinement frames were fabricated with circular cutouts to accomodate the full diameter ingots. The preparation of square tiles with a smaller lateral dimension was not considered advisable, since previous results on ceramics have indicated performance variations for different plate sizes. The circular configuration used provides a roughly equivalent surface area to the standard square tile. In all other respects, target package preparation and testing was identical to the standard ceramic procedure. All ballistic tests were performed at AMTL. A test matrix of the composite targets evaluated is shown in Table III.

Table III. Quantity of Targets Tested

Thickness	As Received		Heat Treated at 1330°C for 20 Hours	
	2.54 cm	5.08 cm	2.54 cm	5.08 cm
5XD (Ti-48 at.% Al + 5 vol% TiB ₂)	1	2	1	2
10XD (Ti-48 at.% Al + 10 vol% TiB ₂)	1	2	1	2
15XD (Ti-48 at.% Al + 15 vol% TiB ₂)	-	2	-	-
7.5XD (Ti-48 at.% Al + 7.5 vol% TiB ₂)	-	-	-	1*

* 3.81 cm in thickness

All data used in this study were taken from shots with a total penetrator yaw under 3 degrees. By means of the known variation of residual penetration with velocity, corrections in measured residual penetration values were made to account for variations in the actual striking velocity. Further details

regarding these procedures are given by Woolsey³⁰. The assumption of a semi-infinite backing plate was checked by examination of the rear surfaces. No loss of flatness was observed for the tested targets. Given this observation, the conditions required for test validity were considered to have been adequately satisfied.

Postmortem fractographic analysis was performed on impacted targets to determine governing dynamic fracture mechanisms in these materials. Representative sections from regions both in and near the penetration interaction zone were taken from the tested targets by EDM to provide samples for metallographic and electron microscopy analysis.

RESULTS AND DISCUSSION

A plot of the residual penetration data for the 5XD, 7.5XD, 10XD and 15XD materials for both as-cast and heat treated conditions is compared to the AD90 alumina reference map in Figure 8. The AD90 map is displayed with upper and lower bound lines representing differences of three standard deviations from the reference curve. Note that the reference map is linear over this range of areal densities. It is apparent that the test results for all of the titanium aluminide materials essentially fall within the normal range of scatter for the AD90 results. Thus, these composites all have the same level of ballistic performance against this penetrator as the reference alumina. This similarity of performance is quite interesting, given the significant differences observed in macro- and microscopic failure morphology from aluminum oxide or other ceramic materials.

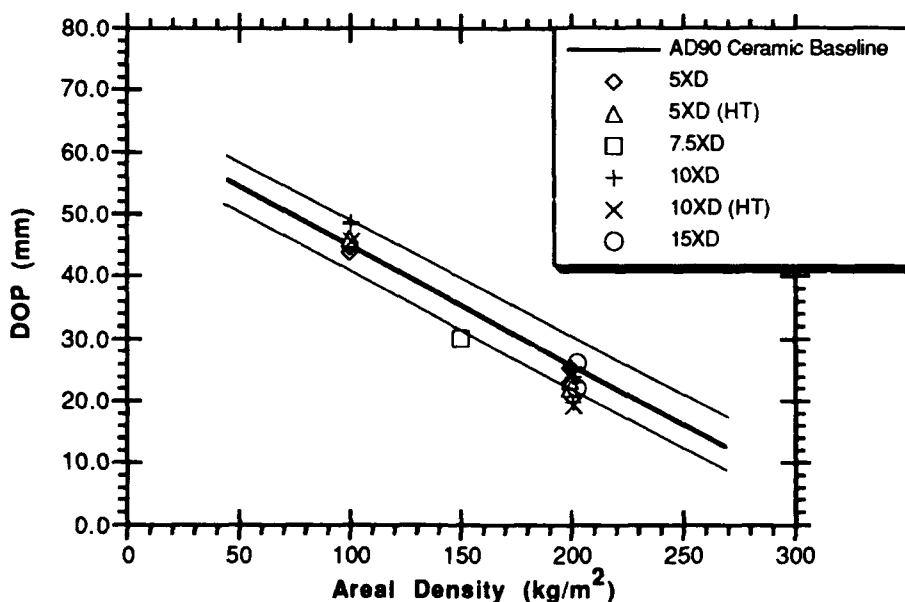


Figure 8. Residual Penetration for Various Titanium Aluminides

Macroscopic examinations of the impacted targets showed a cross-sectional area of the affected material, which roughly conforms to the features of a shallow champagne glass (see Figure

9). For the sake of discussion, region 1 is referred to as the lip, region 2 is referred to as the bowl, region 3 is referred to as the stem, and region 4 is referred to as the base.

Only the 2.54 cm thick as-cast 10XD exhibited the base region. This sample was the only one to separate from its confinement due to the ballistic impact. The 10XD also had the highest hardness among the 2.54 cm thick targets tested.

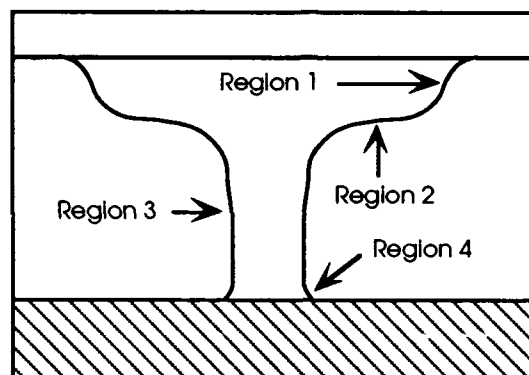


Figure 9. Schematic of Penetration Cavity Cross-Section

Frontal views of a representative series of targets of 5XD, 10XD, and 15XD are shown in Figures 10a-c. The bulk of the material remained basically intact, in contrast to ceramic target plates. Ceramics such as AD90 exhibit extensive damage in the form of radial cracking, ring cracking, and extensive comminution of material in the penetration zone. Some radial cracks have developed in the XD™ titanium aluminide plates. The extent of these in all cases is less than that in a ceramic.

The failure mechanism for target material in the immediate penetration zone is plastic yielding. This is also illustrated by the effective constraint which the TiB_2/TiAl IMC plates provided to the surface of the steel backup block. The degree of plastic flow at the top of the penetration cavity (in the steel) is minimal for all of the tests performed. This is a normal feature of penetration observed between interfaces in metallic materials which undergo ductile yielding, e.g. stacked steel plates. In ceramics, however, extensive fracture at late times due to accumulated damage from the penetration event makes the material in the region of the penetration unable to support shear loading. Hence, it cannot provide an effective constraint to the block face, which responds much like a free surface and flows outward near the penetration cavity.

The diameter of the stem region where the penetrator perforated the target is relatively uniform within each thickness group. The stem diameter of the 2.54 cm targets is noticeably larger than that of the 5.08 cm thick targets. The stem of all the 5.08 cm thick targets is typically filled with a mixture of penetrator and target materials whereas the 2.54 cm targets exhibit a clean hole. The height of the stem appears to be a function of target thickness and micro-hardness. The higher reinforcement loading 10XD composites display a shorter stem compared to the 5XD. Targets with a short stem region also exhibit a deeper and larger diameter bowl area; aspects of this will be discussed later. Microscopic analysis of the stem region revealed significant evidence of local deformation and particle damage (see Figure 11). Apparently, there is sufficient ductility in this material during dynamic compression to prevent rupture due to the impinging penetrator. It is also observed that small cracks as narrow as $50\mu\text{m}$ initiated in this region have become filled and fused with a mixture of penetrator and target material (see Figure 12). Subsequent stresses induced as a result of penetration caused

crack propagation through these filled regions. The role of the penetrator/target material mixture in filling and fusing cracks may have reduced the degree of expulsion of damaged material which would otherwise be expected. More detailed study will be required to obtain a better understanding of this observation.

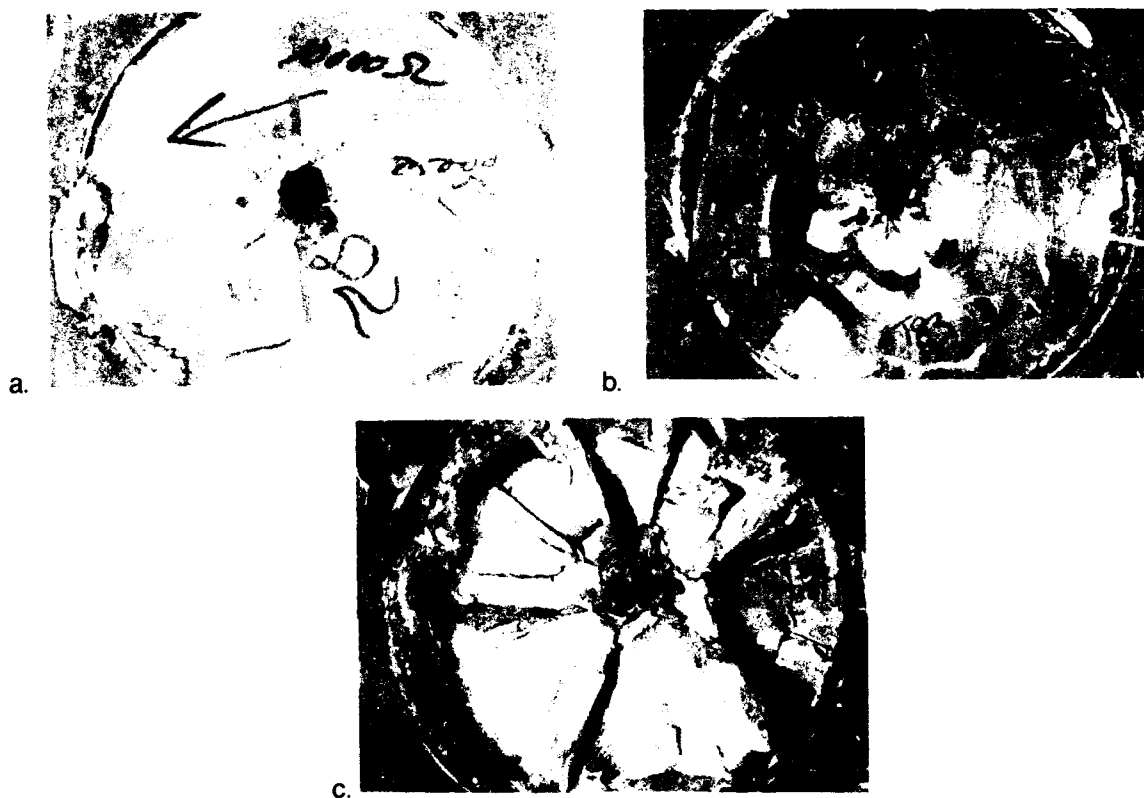


Figure 10. Frontal Surfaces of 5XD(a), 10XD(b), and 15XD(c) Ballistic Targets (Mag 3X)

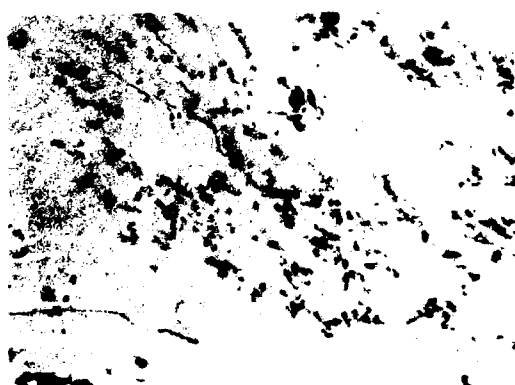


Figure 11. Backscatter Electron Micrograph of Stem Region (Mag 200X)

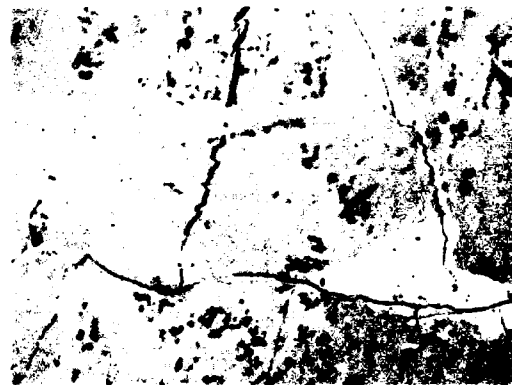


Figure 12. Backscatter Electron Micrograph of W-Fe-Ni Filled Cracks (Mag 200X)

A reverse spallation phenomenon is evident in a majority of the materials tested. Spall is the term applied to failure of a localized region in the target due to the tensile stress, which is created when a compressive shock wave reflects at an interface as a rarefaction wave propagates back into the target. This places material on some plane within the plate above its tensile yield point. It is often observed on the rear surfaces of plates where a ring of material is detached from around the penetration zone and ejected away from the rear of the plate. The position of the spall plane is determined by material and geometric factors of the target material. Preferential failure along the path of microstructural features such as rolling texture is commonly observed. In these tests, the backplate constraint has suppressed the common mode of behavior, but sufficient tensile stresses are generated near the surface to cause detachment and ejection of material in a direction opposite to the penetration event. The extent of this behavior appeared to vary directly with increased particulate loadings (see Figure 10). An excellent example of the operative failure mechanism is observed in the 7.5XD material, where crack propagation is preferential along the direction of particulate alignment developed by the axially symmetric forging process.

Fracture morphologies in the lip and bowl region of the spall zone differ significantly. The fracture surface in the lip region is coarse and jagged, revealing a tortuous path for crack propagation (see Figure 13). The fracture path follows grain boundaries and lamellae directions, and appears to be the result of a relatively low strain rate loading. For 15XD with a high reinforcement loading, the fracture often follows particles in close proximity. In the bowl region, the fracture surface is relatively smooth with extensive secondary crack propagation below the fractured surface (see Figure 14). The general appearance is similar to that for brittle cleavage or fast fracture. It is noteworthy to mention that there is significant particle damage within the fracture surface. The secondary cracks are typically blunted and stopped by various micro-mechanisms. Among the crack blunting mechanisms observed are crack bifurcation within the lamellar structure, and crack blunting resulting from pre-shattered particulates (see Figure 15). Micro-hardness values measured near the fracture surface with significant damaged particles show softening effects. If the particles had been damaged during the initial shock wave, this may soften the material to withstand subsequent secondary shock waves. In the 15XD, cracks propagated as bands with finite thickness (see Figure 16). This observation may relate to the decrease in mean particle spacing with increasing reinforcement loading. Such a feature may be the culprit behind the severe spall exhibited by the 15XD composite.

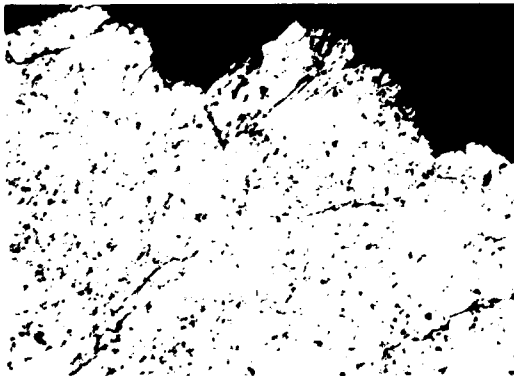


Figure 13. Backscatter Electron Micrograph of a Typical Lip Region (Mag 50X)



Figure 14. Backscatter Electron Micrograph of a Typical Bowl Region (Mag 100X)



Figure 15. Backscatter Electron Micrograph of Crack Bifurcation (Mag 500X)

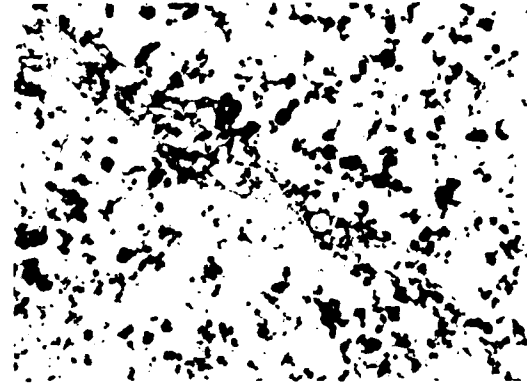


Figure 16. Backscatter Electron Micrograph of a Band Crack in 15XD (Mag 200X)

In summary, the XD™ titanium aluminide composites have comparable ballistic performance to AD90 alumina. The observed level of ballistic resistance was not significantly affected either by microstructure or by reinforcement loading. However, the microstructure played a major role in affecting the spall characteristics of the target. The fracture and crack blunting mechanisms identified are all subject to control by various processing methods. Further work would be needed to tailor the microstructure to minimize spall characteristics. Furthermore, materials with higher reinforcement loading would need to be studied to see the influence of large quantities of ceramic particles on penetration resistance.

CONCLUSIONS

- XD™ titanium aluminide composites exhibit a level of ballistic performance comparable to aluminum oxide ceramics against erosively penetrating long rods.
- The governing failure mode under dynamic impact conditions for XD™ titanium aluminide composites is local plastic yield, which is not comparable to the brittle fracture process exhibited by armor ceramics.
- The observed macroscopic damage evolution in these materials is influenced by microstructural features.
- Micro-fracture mechanisms observed in postmortem analysis include crack blunting due to particulate damage and crack bifurcation in the lamellar structure.

ACKNOWLEDGEMENTS

The authors would like to thank the many people who have contributed to this effort. They would like to particularly thank Dr. C. Robert Crowe of the Naval Research Laboratory and Dr. Steven Kampe of Martin Marietta Laboratories for both providing the materials used to perform this study as well as background information on the characteristics of these materials. Acknowledgement is extended to Professor Ronald Biederman of WPI and Mr. John Nunes for their valuable technical consultations.

The assistance of J. Segalla, J. Loughlin, and R. Muller in performing the ballistic testing, W.F. Cost in collecting test results, and G. Dewing and T. Holmes in sample preparation is also greatly appreciated.

1. G. Pageau, in *Proc. 23rd International SAMPE Technical Conference*, (Oct 1991), pp. 639-649.
2. S.J. Bless et al., Ballistic Impact Behavior of SiC Reinforced Aluminum Alloy Matrix Composites, technical report to be published, (1992).
3. E.Chin, J. Morgan, and F. Tuler, AMTL TR 90-49, (Oct 1990).
4. H. Lipsitt, in *MRS Symp. Proc.*, **39**, (1985), pp. 351-364.
5. S. Ashley, *Mech. Eng.*, **Dec**, 49, (1991).
6. D. Viechnicki et al., Armor Ceramics - 1987, in *Proc. Third TACOM Armor Coordinating Conference*, (Feb 1987).
7. J.M. Brupbacher, L. Christodoulou, D.C. Nagle, U.S. Patent No. 4,710,348, (Dec 1987); U.S. Patent No. 4,772,452, (Sep 1988).
8. D.C. Nagle, J.M. Brupbacher, L. Christodoulou, U.S. Patent No. 4,774,052, (Sep 1988); U.S. Patent No. 4,916,029, (Apr 1990).
9. Y. Kim and D.M. Dimiduk, *JOM*, Aug, 40, (1991).
10. L. Christodoulou, P.A. Parrish, and C.R. Crowe, in *MRS Symp. Proc.*, **120**, (1988), pp. 29-34.
11. D.E. Larsen, S. Kampe, and L. Christodoulou, in *MRS Symp. Proc.*, **194**, (1990), pp. 285-292.
12. E. Robertson and P.L. Martin, in *MRS Symp. Proc.*, **194**, (1990), pp. 233-240.
13. Shih-Chin Huang and D.S. Shih, Microstructure-Property Correlation in TiAl-Base Alloys, Presented at the 1990 Fall TMS Meeting at Detroit, Michigan, Oct. 7-11(1990).
14. Y. Kim and D.M. Dimiduk, *JOM*, Aug, 40, (1991).
15. Shih-Chin Huang, *Met. Trans. A*, **23A**, 375, (1992).
16. D.E. Larsen et al., *Scripta Met. et Mat.*, **24**, 851, (1990).
17. D. Yaziv et al., in *Proc. 9th Intl. Symp. on Ballistics, Part 2*, (1986), pp. 315-319.
18. Z. Rosenberg, S.J. Bless, Y. Yeshurun, and K. Okajima, in *Impact Loading and Dynamic Behavior of Materials*, C.Y. Chiem et al., Eds. (Informationgesellschaft Verlag, 1988), pp. 491-498.
19. Z. Rosenberg and Y. Yeshurun, *Int. J. Impact Eng.*, **3**, 357, (1988).
20. P. Woolsey, S. Mariano, and D. Kokidko, AMTL TR 89-43, (Apr 1989).
21. P. Woolsey, S. Mariano, and D. Kokidko, in *Proc. First TACOM Combat Vehicle Survivability Symp.*, (Mar 1990), pp. 97-106.
22. P. Woolsey, in *Proc. 2nd TACOM Combat Vehicle Survivability Symp.*, (Apr 1991).
23. M. Alme and S.J. Bless, DARPA Report, (1988) [Limited Distribution].
24. M. Alme and S.J. Bless, *ATAC Bullet*, (1989).
25. S.J. Bless, Z. Rosenberg, and B. Yoon, *Int. J. Impact Eng.*, **5**, 165, (1987).
26. R.J. Dowding et al., AMTL TR 90-31, (Apr 1990).
27. V. Hohler and A.J. Stilp, in *Proc. 3rd Intl. Symp. on Ballistics*, (1977).
28. C. E. Anderson Jr. and J. D. Walker, *Int. J. Impact Eng.*, **4**, 481, (1991).
29. E. Chin and R. Biederman, to be published as AMTL TR 92-__, (May 1992).
30. P. Woolsey, in *Proc. 2nd TACOM Combat Vehicle Survivability Symp.*, (Apr 1991).

High Sensitivity Infrared Hot-Electron Transistors

*K. K. Choi, Dr., M. Taysing-Lara, Ms., and W. Chang, Mr.
U. S. Army Electronics Technology and Devices Laboratory
LABCOM, Fort Monmouth, NJ 07703-5601

Recently, thermal imaging has been successfully demonstrated using GaAs/AlGaAs multiple quantum well infrared photodetectors (QWIP). However, the operating temperature tends to be lower than 77 K due to the large dark current existing at this temperature.

We propose to use an additional electron energy filter to reduce the dark current of a QWIP device and hence increase its operating temperature. The resultant device constitutes an infrared hot-electron transistor (IHET). When the infrared detection is dark current limited, we will show that the detectivity of the IHET is increased by a factor of $\alpha_p/\sqrt{\alpha_d}$ in comparison with the original photodetector, and the noise equivalent temperature difference is reduced by a factor of α_p/α_d , where α_p and α_d are the photocurrent and the dark current transfer ratios respectively.

Furthermore, in the voltage coupled readout circuits, we will also show that the IHET structure increases the voltage responsivity of a QWIP by providing an internal photovoltage gain, and the IHET can be dc coupled to a high gain amplifier without a light intensity modulator.

In this paper, I will discuss the basic operating principles and the-state-of-the-art of the QWIP and IHET devices.

A basic QWIP device consists of a number of periods of GaAs quantum wells and $\text{Al}_x\text{Ga}_{1-x}\text{As}$ barriers. The quantum wells are doped to 10^{18} cm^{-3} and the barriers are undoped. The well width is adjusted so that the spacing of the lowest two subbands matches the incoming photon energy. The barrier width is wide enough to suppress the tunneling current from the ground subband. The Al molar ratio x , on the other hand, determines whether the second subband is bounded or extended. The multiple quantum wells are sandwiched between two n doped layers as electrical contacts. For an IHET structure, an additional

$\text{Al}_y\text{Ga}_{1-y}\text{As}$ layer is grown on the top of one of the contacts as an electron energy filter, followed by another n doped layer as the collector.

In the following, we will describe the electrical and optical properties of a typical IHET device. The device is grown by molecular beam epitaxy on a (100) semi-insulating substrate. The first layer grown is a 6000-Å-thick n^+ -GaAs layer as the emitter contact layer. Next, an infrared sensitive QWIP structure is grown. It consists of 50 periods of GaAs wells and $\text{Al}_{0.25}\text{Ga}_{0.75}\text{As}$ barriers. The well width and the barrier width are 40 Å and 200 Å respectively. On the top of the QWIP structure, a thin (300 Å) pseudomorphic $\text{In}_{0.15}\text{Ga}_{0.85}\text{As}$ base layer is grown, followed by a 2000-Å-thick $\text{Al}_{0.25}\text{Ga}_{0.75}\text{As}$ barrier and a 1000-Å-thick GaAs as the collector contact layer. All the layers except the barriers are doped to $1.2 \times 10^{18} \text{ cm}^{-3}$. The band structure is shown in Fig. 1.

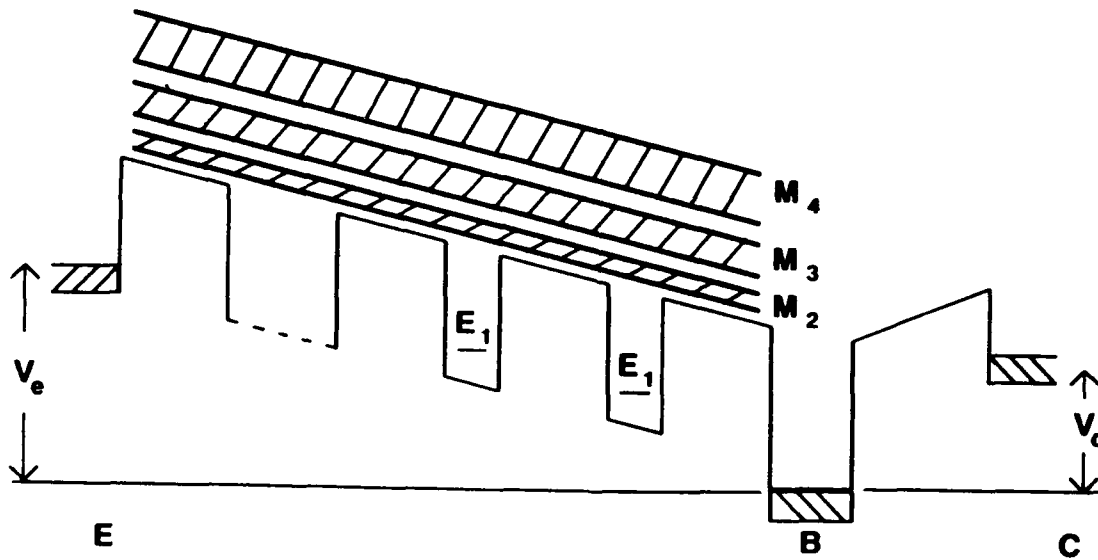


Fig. 1. The band structure of the detector under an emitter bias V_e and a collector bias V_c . M_n are the minibands above the barriers.

Infrared absorption of QWIP structures is initiated by electronic intersubband and transitions from the ground state E_1 to the first excited state E_2 within the

conduction band. Since the dipole moment of the transition is along the superlattice axis, the electric component of the incident light is necessarily nonzero along this direction. The absorption coefficient α_q can be shown to be

$$\alpha_q = \frac{\rho_s}{L_p} \frac{e^2}{2\epsilon_0 m^* c} f \frac{\sin^2 \theta}{n} \frac{\Gamma/2}{(\omega_q - \omega_0)^2 + (\Gamma/2)^2}, \quad (1)$$

where ρ_s is the 2-dimensional electron density per well, L_p is the length per period, f is the oscillator strength, θ is the angle between the direction of light propagation and the superlattice axis, n is the refractive index, and Γ is the absorption linewidth.

For $\rho_s = 4 \times 10^{15} \text{ m}^{-2}$, $L_p = 500 \text{ \AA}$, $f = 0.96$, $\theta = 45^\circ$, $\hbar\omega_0 = 124 \text{ meV}$, $\Delta\lambda/\lambda = 20 \%$, α_{peak} is calculated to be 490 cm^{-1} . If light is allowed to make a double pass through the sample, the internal quantum efficiency η for unpolarized light is then equal to 11 %, estimated from the expression $\eta = 0.5(1 - \exp(-2\alpha l))$, where l is the thickness of the QWIP. The above estimation is done by assuming the infrared light coupled into the device at 45° angle. By using proper diffraction gratings for light coupling, η can be made close to unity.¹ Fig. 2 (a) shows the absorbance of the mentioned QWIP structures as a function of incident photon energy at Brewster's angle. From the peak absorption, we can estimated η at 45° to be 5 %, half of the calculated value. On the other hand, the measured η (= 10 %) of a similar structure is consistent with the calculated value. The second sample has the same nominal well width and doping density but different barrier parameters (width = 300 \AA and Al molar ratio = 0.23), the absorption curve is shown in Fig. 2 (b). The difference in the value of η of the two samples may be due to the difference in the actual doping density and the uncertainty associated with the background subtraction. Note that the two samples have different miniband structures due to different barrier parameters, and the locations of the minibands are also shown in Fig. 2. Despite the differences in the miniband structures, the same peak position at E_2 is observed. This indicates that for single bound state detectors, the optical transition energy is determined by the resonant excited state E_2 of individual wells, which depends only on the well width, but not by the miniband structure. This result shows that the well width is the only adjustable parameter to tune the operating wavelength when the excited state is extended. Also, as shown in Fig. 2 (b), excessive reduction of the barrier height or increasing the barrier width will reduce the detector sensitivity, since the minibands between the top of the barriers and E_2 generate additional thermionic emission current but negligible photocurrent.

On the absorption peak, besides the E_2 peak, there are other modulation structures. These modulation peaks, absent in the normal incidence, match the oscillatory structures in the responsivity curve shown in Fig. 2 (a) measured at temperature equal to 10 K and at 45° light coupling angle. This result indicates that the absorption modulations are due to electron transitions, which can be attributed to the localized states created as a result of coupling between E_1 and the minibands.

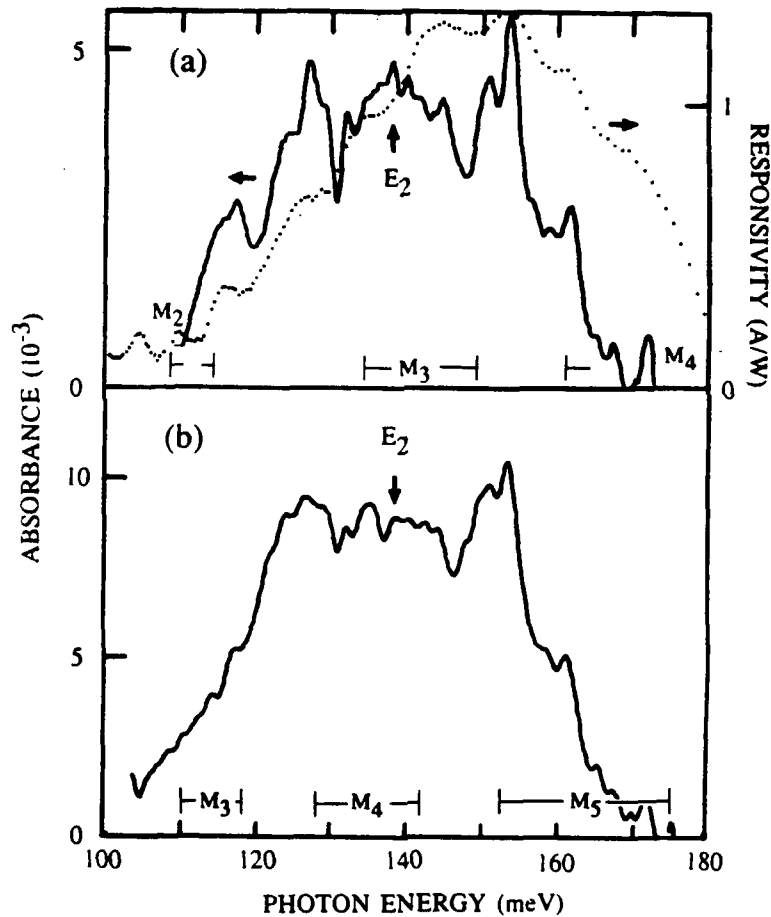


Fig. 2. (a) The absorption spectrum at Brewster's angle and the photocurrent spectrum of a QWIP with well width 40 Å, barrier width 200 Å and Al molar ratio of 0.25. (b) The absorption spectrum a QWIP with well width 40 Å, barrier width 300 Å and Al molar ratio of 0.23.

The dark current characteristics of the detector at 77 K are shown in Fig. 3. For this device, the emitter area is $7.92 \times 10^{-4} \text{ cm}^2$, and the collector area is $2.25 \times 10^{-4} \text{ cm}^2$. The device configuration is shown in the insert of Fig. 3. The emitter current I_e is the current measured at the emitter normalized according to the ratio of the emitter area to the collector area, which also represents the current level of a QWIP device. The electron current injected to the base is separated into two branches according to the energy of the electrons by the energy filter located between the base and the collector. The lower energy electrons are blocked by the barrier and drained through the base, the higher energy electrons overcome the barrier and are collected at the collector and form collector current I_c . The dark current transfer ratio $\alpha (\equiv I_c/I_e)$ as a function of emitter voltage V_e at a constant collector voltage $V_c (= 0 \text{ V})$ depends on the electron energy distribution ρ . In fact, the conducting electron distribution at different temperature can be measured by thermally stimulated hot-electron (TSHE) spectroscopy,² in which r , defined by da/dV_e , is measured as a function

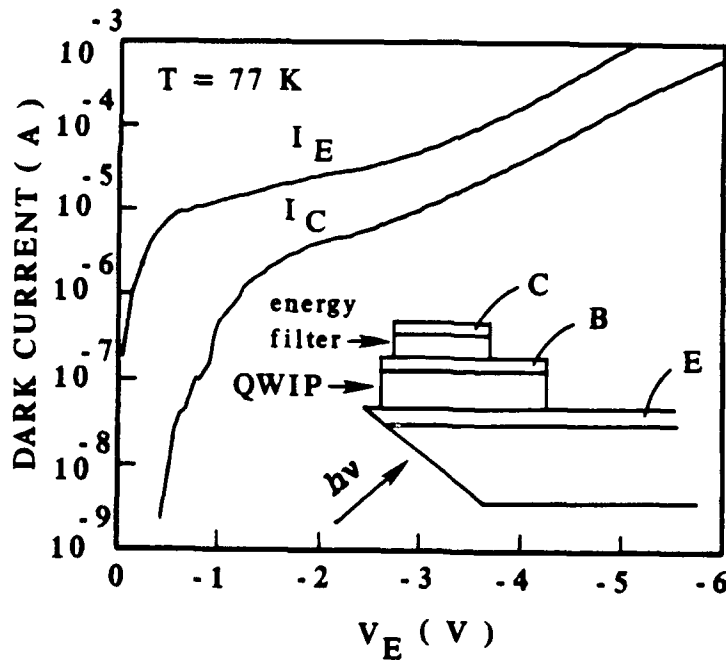


Fig. 3. The emitter dark current I_E and the collector dark current I_C as a function of the emitter bias V_E at 77 K. The insert shows the schematic device configuration.

of V_e at different temperature. Fig. 4 shows the TSHE spectra at temperatures equal to (a) 100 K, (b) 77 K, and (c) 50 K. r can be shown to be proportional to the hot-electron distribution, with the higher energy electrons appearing at the lower V_e . The data show that the peak of the distribution at 77 K is located between E_1 and E_2 , and hence indicate the dominant transport mechanism at this temperature is thermally assisted tunneling.³ At finite temperature, the ground state electrons located at the high energy Fermi distribution tail can be scattered by impurities and tunnel through the quantum barriers in the QWIP structures.

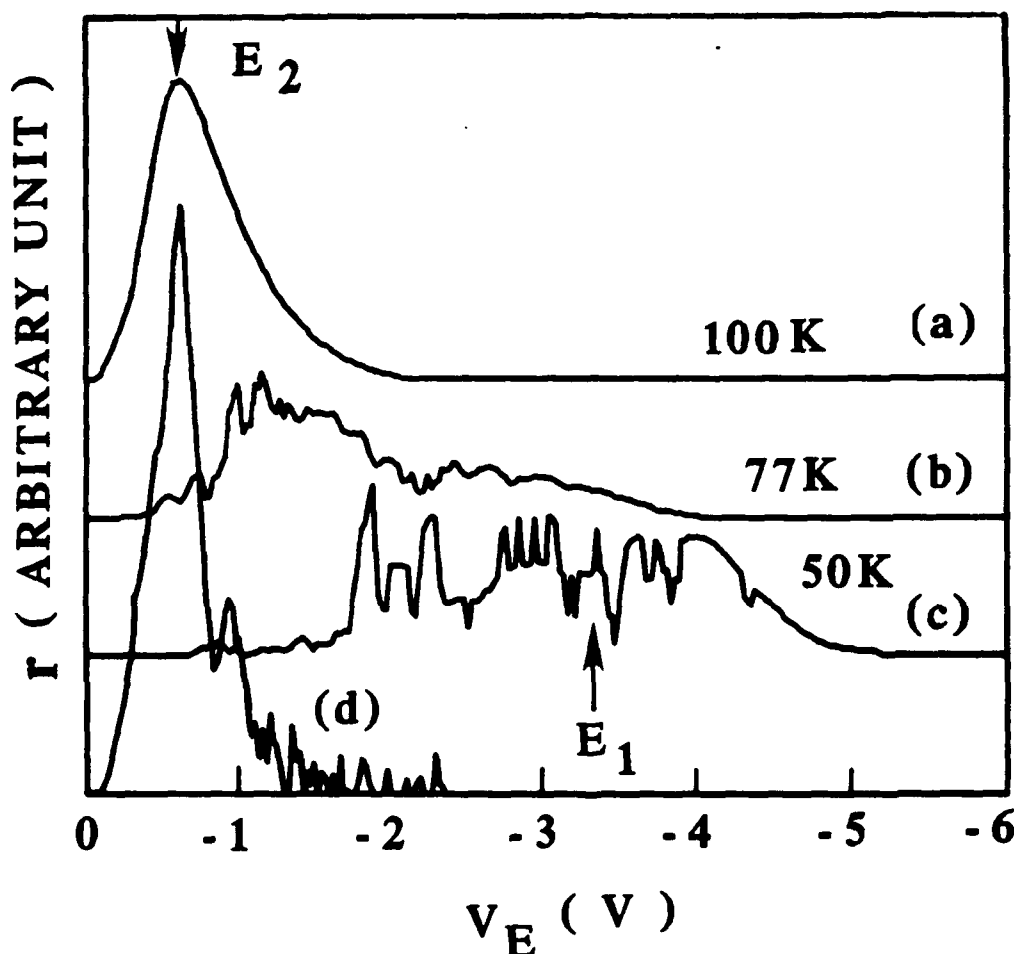


Fig. 4. The parameter r (defined in the text) as a function of V_E at (a) 100 K, (b) 77 K, and (c) 50 K. The parameter r for photocurrent measured at 90 K is indicated by (d). The curves have been shifted vertically for clarity.

The photoresponse of the detector can be characterized using either a CO₂ laser or a glowbar-monochromator. Fig. 5 shows the current responsivity R as a function of V_e . At low V_e , R is directly proportional to V_e . At larger V_e , R becomes constant. After $V_e = -4$ V, R increases again. In order to understand the characteristics of R , it is important to know the photoelectron distribution in the detector. Curve (d) of Fig. 4 is a plot of r_p defined as $d\sigma_p/dV_e$, where σ_p is the ratio of collector photocurrent I_{pc} to the emitter photocurrent I_{pe} normalized to the same area. r_p can also be shown to be directly proportional to the photoelectron distribution. In this measurement, the wavelength of the incident light is 9.25 μm . The result shows that the width of the photoelectron distribution is very narrow. The estimated full width at half maximum (Γ_p) is 17 meV. It is in agreement with the next measurement using different incident photon energy.

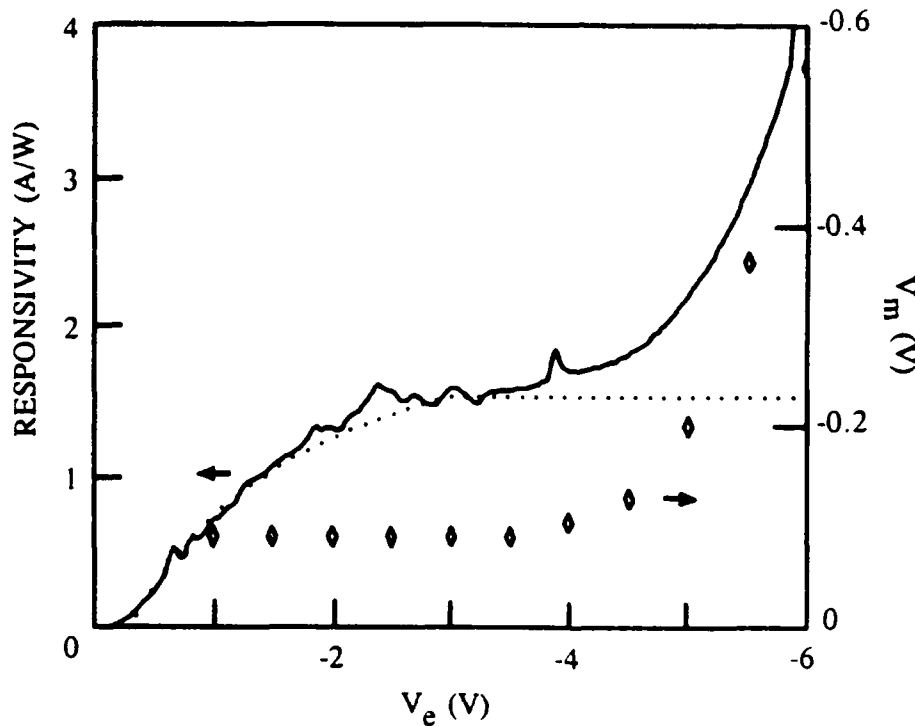


Fig. 5. A plot of experimental (solid curve) and theoretical (dotted curve) DC current responsivity R at photon energy equal to 150 meV as a function of emitter bias V_e . The diamonds are the positions of the photoelectron peak eV_m above the analyzer barrier shown in Fig. 7. R and eV_m show strong correlation.

In the following experiment, the emitter and the collector voltages are fixed at -1.2 V and 0 V respectively. a_p is then measured as a function of incident photon energy. The result is shown in Fig. 6. The solid curve is the fitting to the experimental data assuming that Γ_p is Lorentzian and the photoelectron energy increases with the photon energy. Γ_p is then deduced to be 17 meV at the front boundary of the analyzer as the last experiment. When the detector is under 1000 K radiation, Γ_p is found to be 31 meV, consistent with the absorption width.

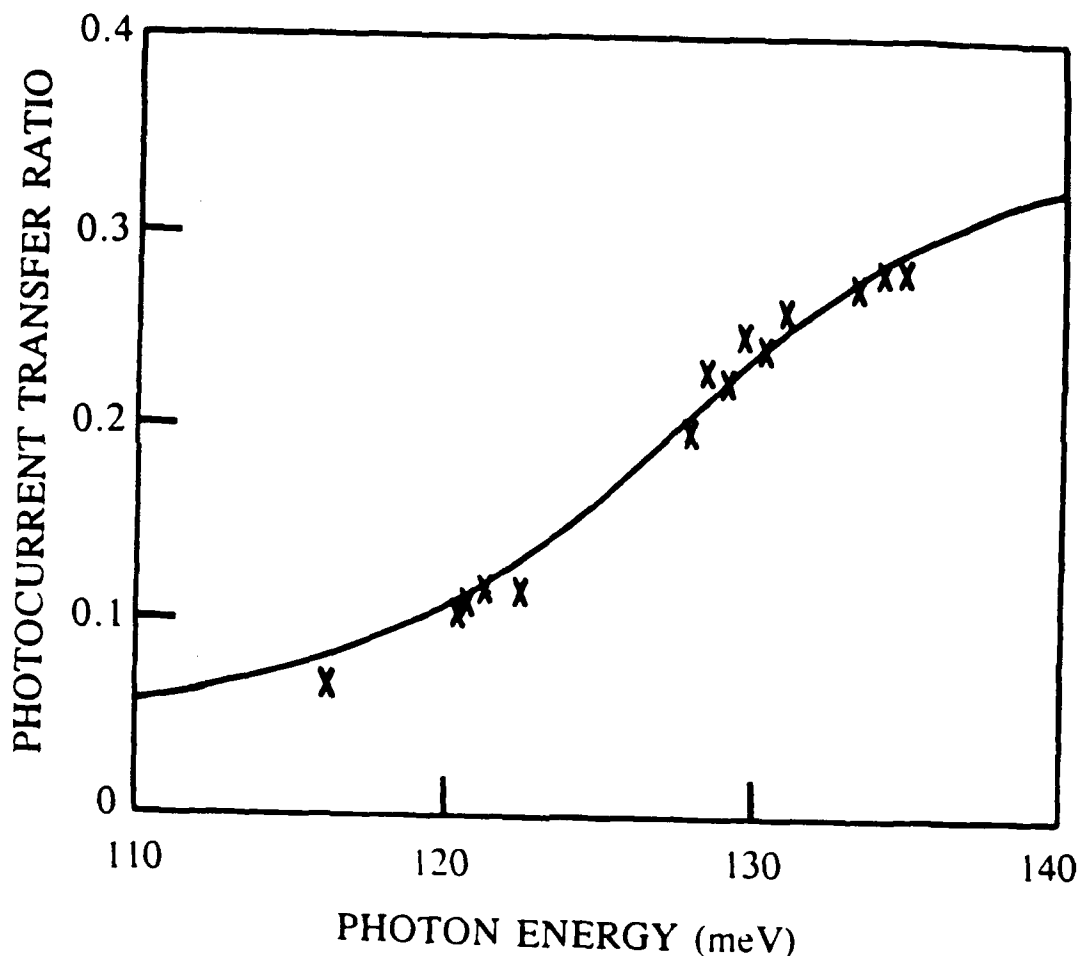


Fig. 6. The crosses are the experimental photocurrent transfer ratio as a function of incident photon energy. The curve is the theoretical fitting assuming that the photoelectron distribution is Lorentzian with a width of 17 meV.

In order to know the photoelectron distribution at different V_e , I_{pc} is now measured as a function of V_c at different V_e . As V_c becomes more negative, the barrier height of the energy filter becomes higher as shown in Fig. 1, and less photocurrent will be collected. The rate of change of I_{pc} as a function of V_c , dI_{pc}/dV_c , is then a measure of ρ . The result is shown in Fig. 7. Γ_ρ is found to be insensitive to V_e and the mean value is 76 meV. However, this apparent Γ_ρ is not the intrinsic width of the photoelectron injection since additional broadening occurs in the thick energy filter. Nevertheless, it does show that the distribution is Lorentzian and the energy of the photoelectrons increases with V_e when V_e is large.

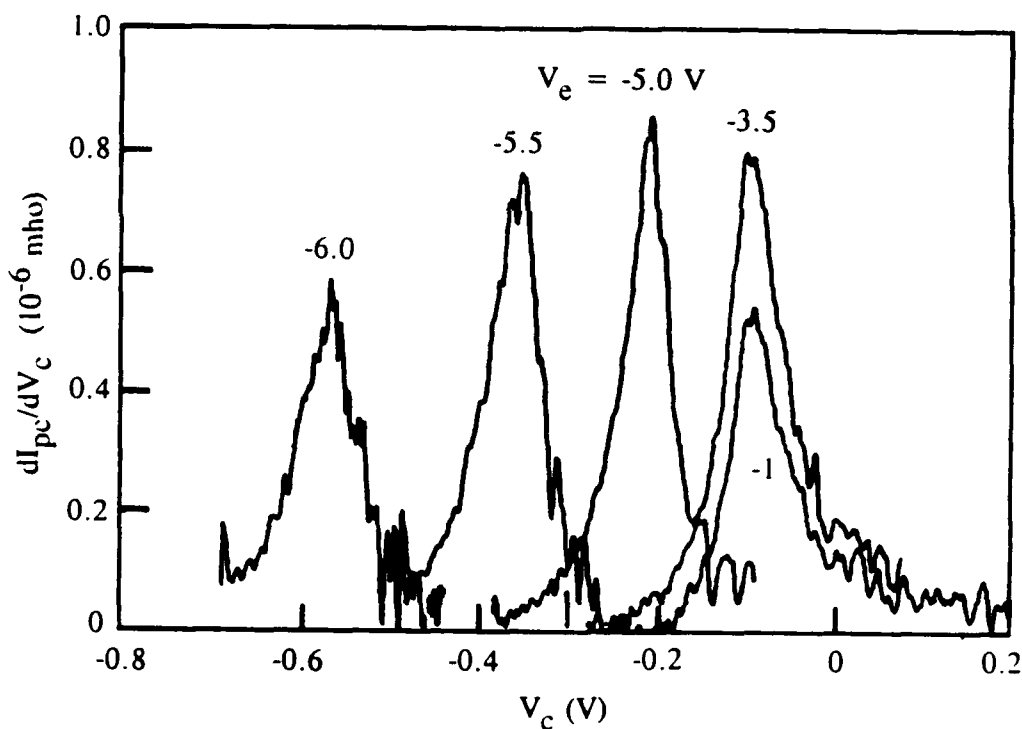


Fig. 7. A plot of dI_{pc}/dV_c as a function of V_c at different V_e at temperature equal to 10 K.

The above results show that the photoelectron distribution is extremely narrow, indicative of the diffusive transport of the photoelectrons; if the transport were ballistic or quasi-ballistic, the expected energy distribution would be much broader. Hence, the standard theory for photoconductor applies for the QWIP devices:

$$R = \frac{e}{h\nu} \eta(V_e) g(V_e) = \frac{e}{h\nu} \frac{\eta(V_e)}{\frac{2v_s}{v_d} + \frac{\tau_{tr}(V_e)}{\tau_o}}, \quad (2)$$

where R is the current responsivity, η is the quantum efficiency, g is the photoconductive gain, v_s and v_d are the contact recombination velocity and the drift velocity respectively, τ_{tr} and τ_o are the transit time and the hot-electron lifetime respectively. Eq. 2 is valid if $v_s \ll v_d$, i.e. hot-electron energy relaxation rate is small at the contacts.

Applying Eq. 2 to the present device, it is necessary to know the specific characteristics of the device. For the present device, high field domain (HFD) formation is observed up to 77 K. Below a critical emitter voltage (V_k) equal to -3 V, an increase in V_e results in an increase in the number of HFDs, but keeping the electric field in the low field regime and the high field regime constant.⁴ Since the major contribution of photocurrent is from the HFDs, the effective η is then proportional to V_e below V_k and constant above V_k . Under a constant electric field below V_k , v_d is also a constant and actually equals to the saturated drift velocity since the electric field (≈ 60 meV per period) is much larger than the critical field (7 meV per period) for maximum drift velocity to occur.⁵ Since v_d is constant, τ_{tr} , being the time to travel across the high field regime, is proportional to V_e below V_k and constant above V_k . If we assume v_s and τ_o to be constant, the DC current responsivity can be fitted as shown in Fig. 5 for $|V_e| < 4$ V with $v_s/v_d = 0.13$, consistent with the assumption, and $\tau_o/\tau_{tr} = 2.5$ at V_k .

For $|V_e| > 4$ V, R is larger than the fitting predicted. In order to understand this discrepancy, it is necessary to examine R at the lower bias. When V_e is small, τ_{tr}/τ_o is less than $2v_s/v_d$, therefore, R is proportional to η and hence V_e . At larger bias, the reverse is true, R is then proportional to $\eta\tau_o/\tau_{tr}$, which is independent on V_e as observed. The rising of R after -4 V then indicates that τ_o increases at large bias. This conclusion is in fact consistent with the increasing

photoelectron peak position at large V_e shown in Fig. 7. As the average energy of the photoelectrons above the quantum barriers is increased, τ_0 is expected to increase. In Fig. 5, we plot the peak position eV_m above the analyzer barrier, which shows strong correlation with R . Note that the contact recombination term we included is crucial to the theory since it explains the voltage dependence at low voltages and determines the ultimate gain achievable at high voltages when τ_0 becomes very large.

In the following, we will estimate the dark current limited detectivity of the detector based on its electrical and optical properties. The detectivity D^* is defined as $R\sqrt{A}/i_n$, where A is the area, R is the current responsivity of unpolarized light at a particular wavelength and i_n is the dark current induced noise current. At 77 K and V_e equal to -0.85 V, i_n at the emitter is measured to be 2.2 pA/ $\sqrt{\text{Hz}}$ (normalized by the square root of the emitter to collector area ratio). Combined with the measured R shown in Fig. 5, D^* is evaluated to be 4.0×10^9 cm $\sqrt{\text{Hz}}/\text{W}$ at 8.2 μm . Likewise, D^* is equal to 5.5×10^9 cm $\sqrt{\text{Hz}}/\text{W}$ and 6.0×10^9 cm $\sqrt{\text{Hz}}/\text{W}$ at $V_e = -1.33$ V and -1.8 V respectively. Note that the value of the gain deduced from the noise current at -0.85 V is 0.66, from the relation $i_n = (4eg\mu_d)^{1/2}$, differing from the value of 3.0 we deduced from Eq. 2. The discrepancy is due to the fact that the gain that determines the noise current is the average gain of the entire QWIP, including the high field regime and the low field regime, while the gain determined in Eq. 2 is the average gain at the high field regime only.

For the D^* of an IHET, the current responsivity R is reduced by a factor of α_p compared with an QWIP, but at the same time, the dark current is also reduced by a factor of α_d where α_p and α_d are the photocurrent transfer ratio and the dark current transfer ratio respectively. As a result, D^* of an IHET is increased by a factor of $\alpha_p/\sqrt{\alpha_d}$ in comparison with the original QWIP. Since the energy of the photocurrent is higher than the energy of the dark current at 77 K as indicated in Fig. 4, α_p can be much larger than α_d at certain V_e . For example, at -0.85 V, α_p is 0.33 which is 36 times larger than $\alpha_d (=9.1 \times 10^{-3})$. Hence, the D^* of the IHET is increased by a factor of 3.5 at this bias, and is equal to 1.4×10^{10} cm $\sqrt{\text{Hz}}/\text{W}$. At lower temperatures, it is 3.0×10^{10} , 1.1×10^{11} , and 7×10^{11} cm $\sqrt{\text{Hz}}/\text{W}$ at 70 K, 60 K, and 50 K respectively, higher than the corresponding QWIP devices.

Note that the D^* of an IHET can exceed the estimation given by Kinch and Yariv⁶ for the QWIPs. In their analysis, the thermal electrons include all the electrons with energy higher than the barrier height (= 190 meV in the present case), while an IHET is capable of discriminating all the electrons up to the energy (= 215 meV) corresponding to the cutoff wavelength. This difference alone reduces the dark current by a factor of 43, and hence increases the estimated D^* .

by a factor of 6.6 at 77K.

Besides the D^* , another figure of merit of a detector array is the noise equivalent temperature difference (NE Δ T). When the IHETs are used in a detector array with charge coupled device (CCD) readout circuit, the integration time per frame can be made longer by a factor of $1/\alpha_d$ than in the case with the QWIP, due to the reduced dark current. As a result, NE Δ T of the detector array is a factor of α_p/α_d ($= 36$ in the present case) smaller when the IHET structure is used due to the increased D^* and the longer integration time.

In the last section, we have shown the advantages of IHET in the charge coupled detector arrays. In the following, we would like to show that IHET is also advantageous in voltage coupled readout circuit. In scanning detector arrays, high gain operational amplifiers are often used instead of CCDs due to their lower cost. In this case, photovoltaic detectors such as photodiodes are much more favorable than photoconductors, mainly because being under a large dc bias, photoconductors often need ac coupling and hence require some type of light intensity modulation during the operation. In the present voltage coupled readout circuit, we will show that the IHET structure can also be used to improve the characteristic of a QWIP in two significant areas, specifically, the IHET provides an internal photovoltage gain which increases the voltage responsivity of a QWIP, and the IHET can be dc coupled to a high gain amplifier, thereby eliminating the requirement of light modulation.

In order to understand the amplification mechanism of an IHET, let us first examine the photovoltage (V_p^c) measured from a QWIP which is given by

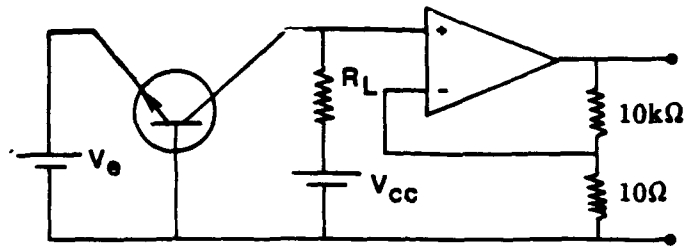
$$V_p^c = i_p \left(\frac{1}{r} + \frac{1}{R_L} \right)^{-1}, \quad (3)$$

where i_p is the emitter photocurrent, r is the internal dynamic resistance of the QWIP, and R_L is the external loading resistance. If the same QWIP is incorporated into the emitter of an IHET structure, the photovoltage (V_p^t) measured across R_L , according to the standard transistor theory, is easily shown to be

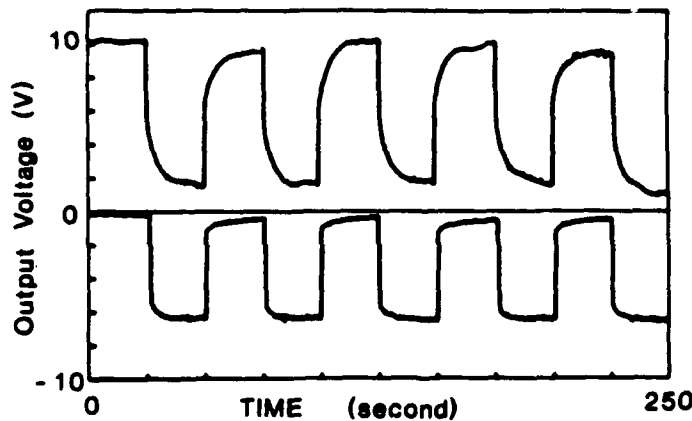
$$V_p^t = \alpha_p i_p \left(\frac{1}{R} + \frac{1}{R_L} \right)^{-1}, \quad (4)$$

where $1/R$ is the output admittance with input open. In the present case, $R = 4.6 \text{ M}\Omega$ is much larger than $r = 69 \text{ k}\Omega$ at 77 K and $V_e = -0.85 \text{ V}$. If we assume the loading resistance in both the QWIP and IHET is equal to r and R respectively, then the photovoltage gain will be equal to $\alpha_p R/r$ which is 22. However, in order

to achieve this high gain, the loading resistance in the IHET case is very large. One has to apply a large positive collector bias V_{cc} as shown in Fig. 8 to balance out the voltage drop across R_L due to the ambient dark current flow. If we limit V_{cc} to be less than 0.2 V, the largest R_L one can adopted is 2 M Ω in order to keep the operating point Q above the threshold voltage V_{th} (≈ -0.15 V). In this figure, V_c is the voltage between the collector and the base. The experimental value for the photovoltage gain is 9.7 at $V_e = -0.85$ V when R_L of the IHET is 2 M Ω and V_{cc} is 0.2 V. The voltage responsivity is measured to be 130 kV/W.



(a)



(b)

Fig. 8. The schematic diagram of an IHET dc coupled to a voltage amplifier, (b) the dc output voltage of the amplifier at 11 K (lower curve) and at 77 K (upper curve). The 77 K data have been shifted up by 10 V for clarity.

For a given R_L , one can adjust V_{cc} such that the operating point Q is on the axis where V_c is zero. Under this condition, V_c is nonzero only when there is a photosignal, similar to the photovoltaic operation. The IHET can then be directly coupled to a high gain voltage amplifier without saturating the amplifier output. Fig. 8 (a) shows an amplifier circuit with a voltage gain of 1000. Fig. 8 (b) shows the output voltage of the IHET with $R_L = 100 \text{ k}\Omega$ and $V_{cc} = 3 \text{ mV}$ at $T = 10 \text{ K}$ and 8 mV at 77 K . These data are obtained by manually opening and closing an optical shutter in a regular time interval, with the initial shutter position close. Note that the ambient output voltage without light is very close to zero volts, showing the flexibility of the transistor in dc coupling. The measured noise is temperature independent in the present case, which indicates that the light detection process is amplifier noise limited. Finally, it is interesting to note that the schematic diagram in Fig. 8 (a) may be viewed as a two stage amplifier, since the light signal detected at the emitter is first amplified by the transistor and subsequently by the operational amplifier.

In summary, we have elucidated the basic operating principle of QWIP and IHET devices. In general, the IHET improves the characteristics of an QWIP significantly. For the present IHET, it increases the detectivity of a QWIP by a factor of two, reduces the noise equivalent temperature difference by a factor of 36, increases the photovoltage response by a factor of 9.7, and provides quasi-photovoltaic operation. The detectivity of the present detector is $1.4 \times 10^{10} \text{ cm}\sqrt{\text{Hz/W}}$ at 77 K . With further optimization of device parameters such as doping density and the structure of the QWIP, the structure of the energy filter, and the scheme of optical coupling, a broadband $10 \text{ }\mu\text{m}$ IHET with D^* close to $10^{11} \text{ cm}\sqrt{\text{Hz/W}}$ should be achievable at 77 K .

Reference

1. J. Y. Andersson and L. Lundqvist, Appl. Phys. Lett. **59**, 857 (1991).
2. K. K. Choi et al., Appl. Phys. Lett. **57**, 76 (1990).
3. E. Pelves et al., J. Appl. Phys. **66**, 5656 (1989).
4. K. K. Choi et al., Phys. Rev. B **35**, 4172 (1987).
5. X. L. Lei et al., Phys. Rev. Lett. **66**, 3277 (1991).
6. M. A. Kinch and A. Yariv, Appl. Phys. Lett. **55**, 2093 (1989).

CHRISTENSEN, CLARKE, GRISHAM, TANTON

Optical Mapping of HgCdTe Detector Electronic Properties

Charles R. Christensen, Dr.* , Frederick W. Clarke, Mr., and
John A. Grisham, Mr.

Weapons Sciences Directorate, AMSMI-RD-WS-CM
Research, Development, and Engineering Center
U. S. Army Missile Command, Redstone Arsenal, AL 35898-5248

and

George A. Tanton, Dr.
Teledyne Brown Engineering
300 Sparkman Drive, Huntsville, AL 35807-7007

1.0 INTRODUCTION

The primary objective of this work was to assess the Faraday rotation technique as a method to characterize semiconductor wafers for improved efficiency in focal plane array production that could support strategic and tactical programs. The approach has been to work closely with material growers to compare electronic properties measured in this laboratory using magneto-optical techniques to those measured on the same samples using standard methods. Work in this laboratory on the application of Faraday rotation (FR) technique to determine free carrier concentrations in mercury cadmium telluride (MCT) and other semiconductor materials has been previously reported. Recent upgrades in the laboratory technique and hardware have resulted in new temperature- and wavelength-dependent data on MCT. A goal in analyzing this new data was to assess the feasibility of obtaining other information, e.g., mobility, in addition to free carrier wafer characterizations from the FR technique, resulting in a more complete wafer characterization. Mobilities were calculated from experimentally measured FR and absorption data on materials previously characterized by the Hall technique.

If the incident photon energy is not too near the semiconductor bandgap energy the Faraday rotation, δ , can be described by¹

$$\delta = a\lambda^{-2} + b\lambda^2 \quad (1)$$

The first term of Eq. (1) is the interband Faraday rotation and is not a function of the free carrier concentration. The interband term is large for wavelengths near the absorption edge. The second term is the free carrier contribution to the Faraday rotation and is related to the magnetic field strength, B , and carrier effective mass m^* by

$$b = \frac{Ne^3 \ell B}{2\pi c^4 m^{*2} n} \quad (2)$$

where N is the free carrier concentration, e is the electron charge, ℓ is the material thickness, c is the speed of light, and n is the refractive index. Both terms in Eq. (1) contribute significantly to Faraday rotation in HgCdTe at CO₂ laser wavelengths.

2.0 EXPERIMENTAL

The experimental apparatus, shown schematically in Figure 1, is similar to that described previously². In the work reported here the CO₂ laser was tunable from 9.2 to 10.8 μm and the output was monitored by the power meter and spectrum analyzer addressed by a movable mirror. Nearly simultaneous measurement of rotation and absorption was made at each point on the sample by using two lock-in amplifiers, one tuned to the chopper frequency and the other tuned to the magnet modulation frequency. A calibration factor in millivolts per degree of Faraday rotation for the system was determined by comparing the output signal with the polarization rotation induced by a dc magnetic field. The samples were mounted on a cold finger in a dewar having 2.0 mm thick ZnSe windows. The sample was held rigidly to maintain point-to-point orientation throughout the experiment. Faraday rotation measurements at 3 temperatures were made at 15 points across a 2 mm \times 4 mm area of the sample at 7 CO₂ wavelengths between 10.74 μm and 9.2 μm . To compensate for random laser drift in intensity, an average over three runs was used.

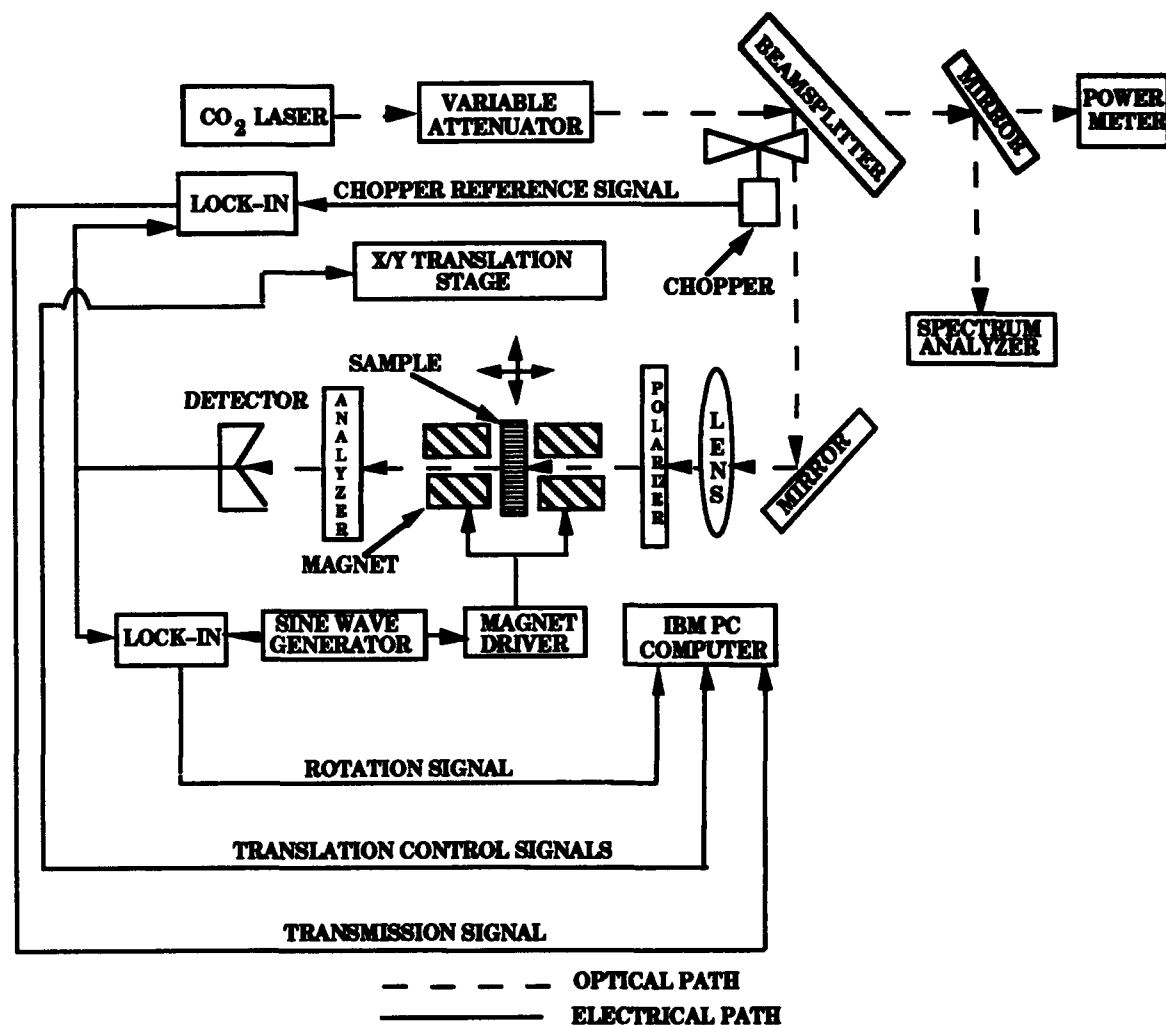


Figure 1. MICOM Faraday Rotation Mapper

Several samples of $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ were obtained from the Central Research Laboratories of Texas Instruments, Inc., Dallas, Texas. Nominal x values were determined by cutoff wavelengths measured by FTIR spectroscopy; free carrier concentrations and mobilities were determined by Hall effect measurements. A sample of nominal x value 0.274, transparent to CO_2 laser wavelengths at 77°K, was chosen. The sample had a thickness of 1.08 mm, and average free carrier (electron) concentrations of $5.80 \times 10^{15} \text{ cm}^{-3}$ and $2.38 \times 10^{14} \text{ cm}^{-3}$ at 300°K and 77°K, respectively, with corresponding mobilities of $6.16 \times 10^3 \text{ cm}^2 (\text{Vs})^{-1}$ and $5.66 \times 10^3 \text{ cm}^2 (\text{Vs})^{-1}$.

3.0 DOPANT AND FARADAY ROTATION COMPONENT MAPPING IN HgCdTe

As indicated in Eq. (1) above, the FR signal is composed of two main components, the interband component, and that due to the free carriers in the material, both intrinsic and extrinsic (dopant carriers). The purpose of measuring Faraday rotation at different wavelengths and different temperatures is to separate the interband and the intrinsic and dopant carrier components. The components were separated using a method described by Ahrenkiel³. Faraday rotation measurements for each point were plotted versus wavelength at a particular temperature and a least squares line was fitted to the points. Two arbitrary points on the least squares line (say at 9.0 μm and 11.0 μm) were used in Eq. (1) to solve for a , the interband coefficient, and b , the free carrier coefficient. This method assumes that Eq. (1) is linear with respect to wavelength over the narrow range of CO₂ laser wavelengths. Least squares plots, one for each point for each temperature for a total of 45 plots, were found to be nearly linear for these experiments. Figure 2 is a mapping of the free carrier component of the FR at 15 points across the sample at 77°K. The calculated negative rotations from the least squares plots are slightly higher than the experimental values. This is due to a slight negative rotation from the dewar zinc selenide windows. The offset rotation is independent of sample temperature since the windows are not cooled. According to the calculated curves of intrinsic carrier concentration versus x of Long and Schmit⁴, the intrinsic carrier concentration of this sample ($x = 0.274$) at 77°K is on the order of 10^6 less than that at room temperature. With the intrinsic carriers thus frozen out, the only source of free carrier rotation is that due to the dopant carriers, whose numbers are more or less temperature-independent. Thus the free carrier components of the FR at 77°K in this sample are due entirely to the dopant and this figure is a mapping of the spatial doping profile across the sample. As can be seen from the figure, the free carrier FR at 77°K is negative at all points. Since electrons give rise to negative FR (note that the interband FR is positive), the sample is everywhere n-type with variations in dopant concentration as shown in the figure. The free carrier components at 200°K and 300°K are due to the combined dopant carriers (whose numbers at 200°K and 300°K are the same as they were at 77°K) and the intrinsic carriers at those temperatures. Mappings of these components are mappings of the total free carrier population at that temperature. Figure 3 shows such a map at 300°K. The average carrier concentrations measured on this wafer by Texas Instruments were used to plot these figures. It will be necessary to calibrate our system with a standard reference sample to determine absolute carrier concentrations from Eqs. (1) and (2).

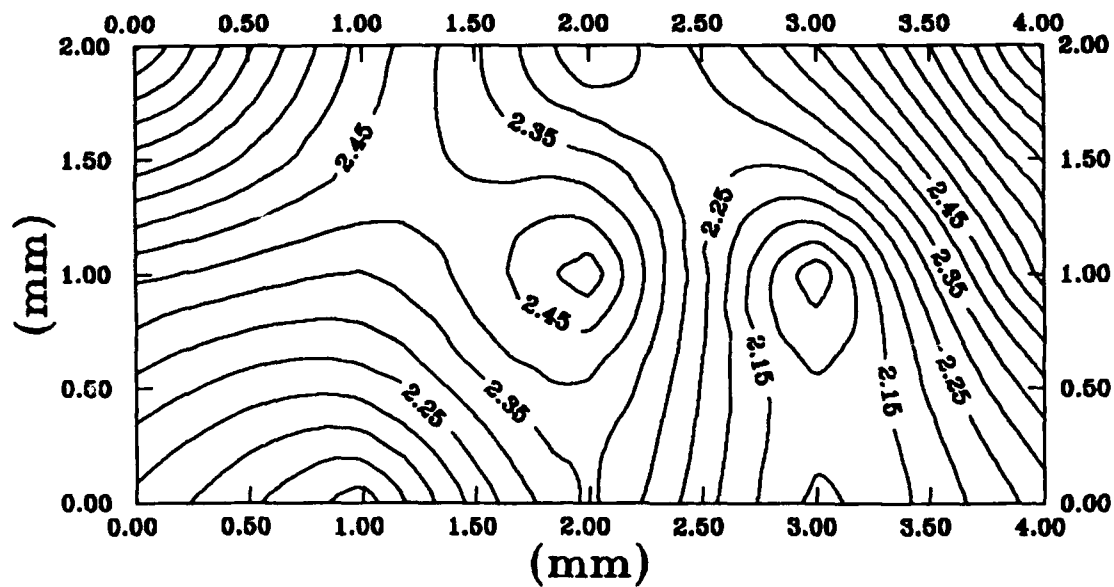


Figure 2. Doping Profile at $77^{\circ}\text{K} \times 10^{14} \text{ (cm}^{-3}\text{)}$

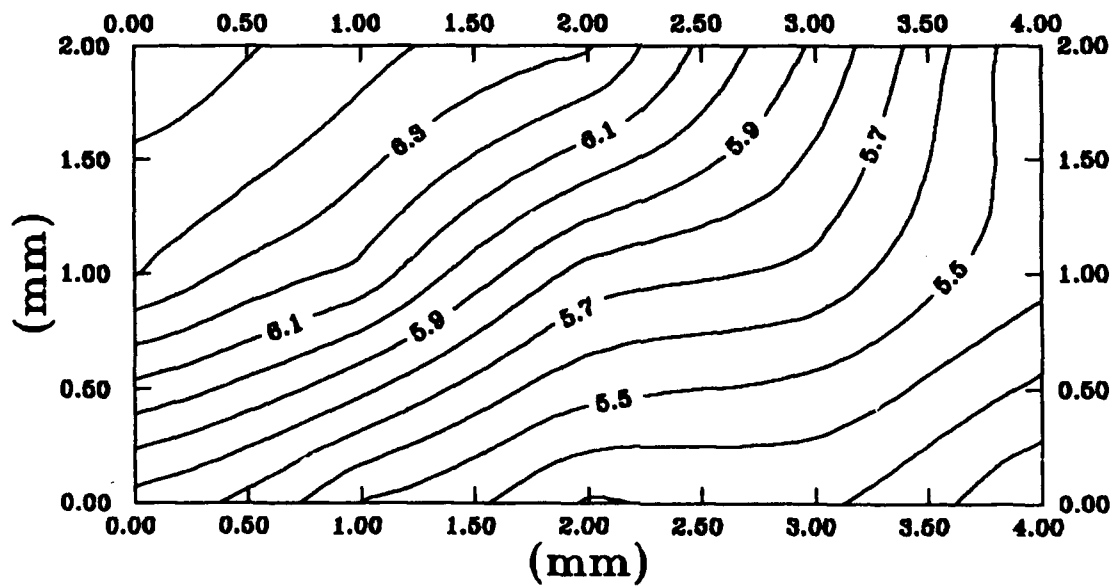


Figure 3. Total Free Carriers at $300^{\circ}\text{K} \times 10^{15} \text{ (cm}^{-3}\text{)}$

Separation of the intrinsic and dopant FR can be accomplished at a particular temperature by scaling the FR due to dopant carriers at 77°K by the ratio of the square of electron effective masses at 77°K and at the desired temperature according to Eq. (2). The intrinsic component is obtained by subtracting the scaled dopant FR from the total free carrier component. The intrinsic component at each point at 300°K is mapped in Figure 4. Since the intrinsic carrier concentration is proportional to the x concentration according to the curves of Long and Schmit⁴, this map shows the spatial variation of x across the sample. Since the intrinsic carrier concentration varies inversely with x , those areas of increased FR are areas of *smaller* x . Note that since the electron effective mass also varies with x , scaling the dopant FR component at each point would require that x be known at each point. However, Fig. 4 shows that the intrinsic FR for the area of the sample measured in this experiment is reasonably constant within $\pm 15\%$ of the mean. Intrinsic carriers (and their associated FR) drop off considerably³ with x . For example, a 10% variation in x in the region of $x = 0.274$ causes about a $\pm 85\%$ variation in carrier concentration at 300°K. The FR is further affected by at least $\pm 10\%$ by the differences in effective mass squared. This would cause an over 100% variation in FR due to a 10% variation in x . Thus a $\pm 15\%$ variation in FR indicates a fairly small variation in x . This justifies the use of a single value of x (the nominal value of 0.274) as an approximation to scale the dopant rotations to produce the map in Figure 4.

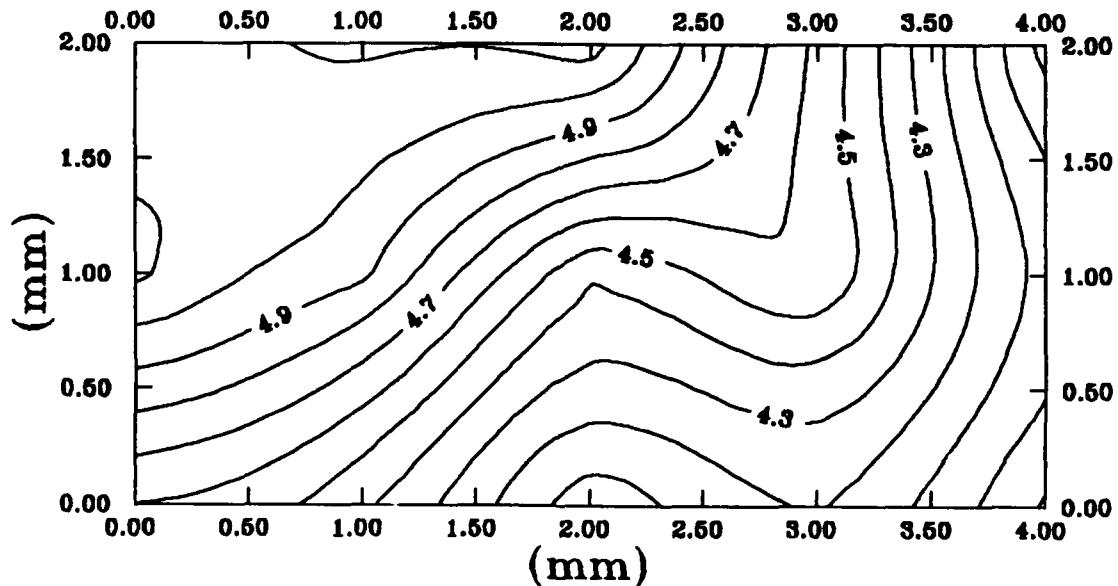


Figure 4. Intrinsic Carriers at 300°K $\times 10^{15}(\text{cm}^{-3})$

4.0 MOBILITY INFORMATION FROM FARADAY ROTATION AND TRANSMITTANCE

Faraday rotation and absorption data can be used together, in principle, to yield mobility information.^{5,6} Mobility is calculated in this section from data taken on two similar wafers to demonstrate that the technique yields reasonable results.

A Faraday rotation angle of 3 degrees determined previously for a $\text{Hg}_{.78}\text{Cd}_{.22}\text{Te}$ wafer furnished by CNVEO, U. S. Army CECOM, with a calculated intrinsic carrier concentration of 2×10^{16} at 300°K ⁷ was used. An absorption coefficient, $\alpha = 7.3 \text{ cm}^{-1}$ was calculated using Eq. (3) and a transmittance value, $\tau = 0.25$, that was determined experimentally for an arbitrary location on a similar $\text{Hg}_{.782}\text{Cd}_{.218}\text{Te}$ wafer using the current Faraday rotation set up. Values of $x = .218$ and $N = 2.70 \times 10^{16}$ were supplied by Texas Instruments, Inc. for this wafer.

Transmittance τ is given by

$$\tau = (1-\rho)^2 e^{-\alpha \ell} / 1-\rho^2 e^{-2\alpha \ell} \quad (3)$$

where $\rho = [(n-1)/(n+1)]^2$.

In Eq. (3), α is the free carrier absorption coefficient and ρ is the reflectivity. The index of refraction was taken to be⁹, $n = 3.3$ and was assumed to remain constant. The mobility μ was calculated from⁶

$$\mu = 2\theta / \alpha \ell B \quad (4)$$

for a given temperature, where θ = Faraday rotation angle, B and ℓ are the magnetic field strength and wafer thickness, respectively.

From Eq. (4)

$$\begin{aligned} \mu &= 0.105/7.3 \left[\text{cm}^{-1} \right] \times 0.0864 \left[\text{cm} \right] \times 2400 \left[\text{G} \right] \\ &= 0.69 \left[\text{m}^2 \right] / \left[\text{Weber} \right] = 0.69 \times 10^4 \left[\text{cm}^2 / \text{Volt sec} \right] \end{aligned}$$

From data obtained with the Faraday rotation technique we calculated a mobility value of $0.69 \times 10^4 \left[\text{cm}^2 / \text{Volt sec} \right]$, as shown above, which falls within

the range of $(0.547 \text{ to } 1.28) \times 10^4$, given for samples at $T = 296^\circ\text{K}$ with $N = (1.21 \text{ to } 3.11) \times 10^{16}$ from Texas Instruments, Inc.

Using the method described above, mobilities, μ , were calculated for several places on the mercury cadmium telluride wafer described in Section 2.0 using transmittance data obtained at $T = 300^\circ\text{K}$ from the current Faraday rotation setup. The detector signal levels when the wafer was in the beam and with the wafer removed were used to calculate α , and is displayed in Fig. 5. When calculating an absorption coefficient, only reflection from the front surface of the sample was taken into account. Figure 6 is a map showing the mobility variation of the sample. The average mobility for the mapped area was calculated to be $3120 \text{ [cm}^2/\text{Volt sec]}$ which is approximately half of the $6160 \text{ [cm}^2/\text{Volt sec]}$ value reported by Texas Instruments for this wafer.

5.0 CONCLUSIONS

Although free carrier absorption dominates in most semiconductors at CO_2 wavelengths, there is still a significant crystal lattice component in the absorption coefficients measured in this experiment. This partially explains the low electron mobility values obtained at room temperature and would affect low temperature measurements even more. Since mobility is severely affected by crystal defects, more accurate determination of mobility across the sample could be used to map some kinds of crystal defects. This would necessitate formulation of a method for separation of the free carrier and lattice components from the absorption measurements available from the present setup. Once this is accomplished, mobility at each point on a sample could be obtained at several temperatures and compared with reference high quality crystals. Point-to-point differences could then be mapped as a quantitative measurement of various crystal defects, effecting mobility and presumably infrared detector performance. Some judgement as to the type of defect could be possible by comparison of these differences at different temperatures, since different scattering mechanisms are present at different temperatures.

We conclude that the Faraday rotation technique can provide rotation and transmission data from which total free carrier, dopant, and mobility information can be obtained, resulting in a more complete wafer characterization. Using the approach just described, maps of total free carrier, intrinsic carrier, doping profile, and mobility variation over a wafer were generated from the FR technique.

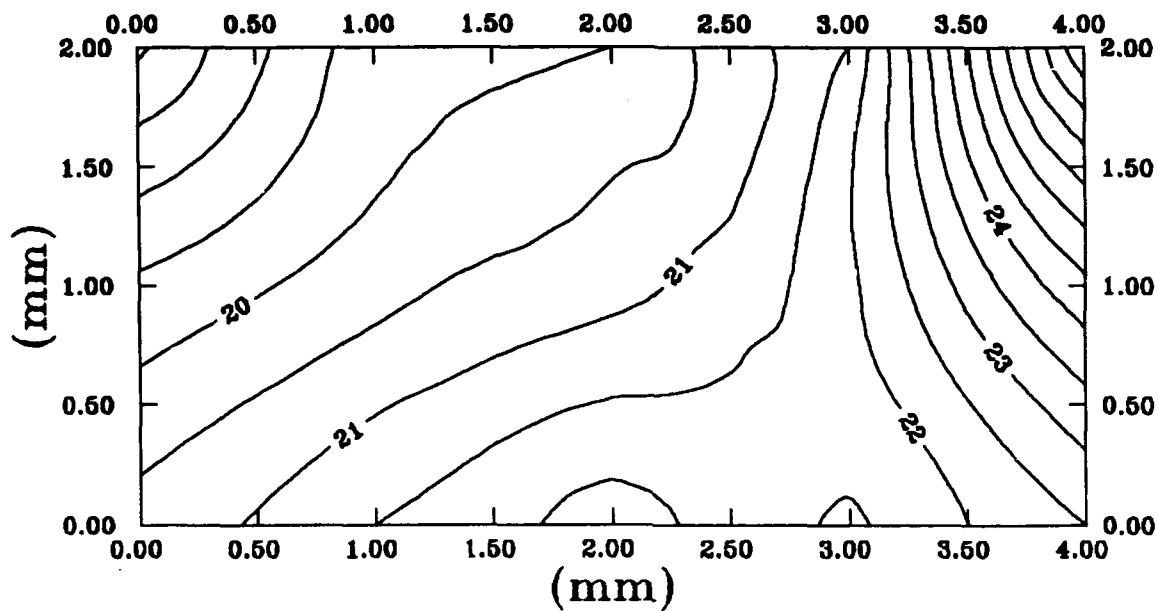


Figure 5. Absorption Coefficient at 300°K (cm^{-1})

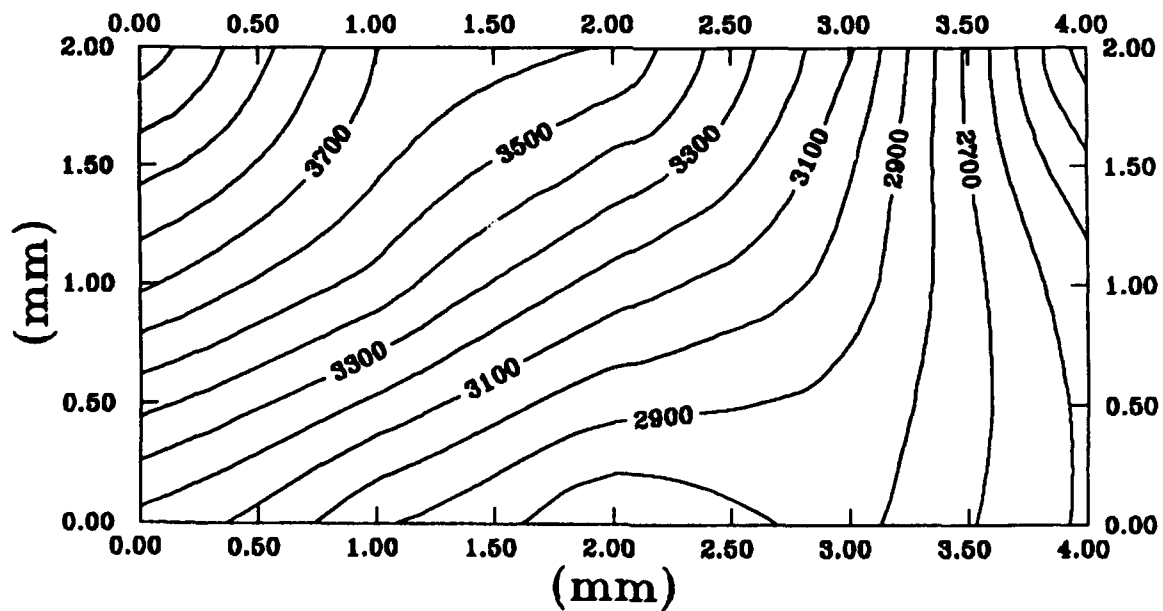


Figure 6. Mobility at 300°K ($\text{cm}^2/\text{Volt Sec}$)

CHRISTENSEN, CLARKE, GRISHAM, TANTON

Transfer of FR technology through close interaction with the material grower/user community is expected to make this a useful evaluation tool for detector production and material development. In addition to MCT, application of the FR technique to other materials, e.g., GaAs, InSb, CdTe, and quantum well structures should be considered.

ACKNOWLEDGEMENTS

The characterized HgCdTe wafers were provided by A. J. Syllaos of the Central Research Laboratories, Texas Instruments, Inc., Dallas, TX. The software used in data collection and processing was written by John Stensby of the Electrical and Computer Engineering Department of the University of Alabama in Huntsville, AL, and by T. C. Leslie of this Directorate. This work was supported through the Manufacturing Science Program of the Systems Engineering and Production Directorate's Manufacturing Technology Division and by the Weapons Sciences Directorate, RD&E Center, U. S. Army Missile Command.

REFERENCES

1. M. Balkanski and J. J. Hopfield, *Phys. Stat. Sol.* **2**, 623 (1962).
2. John L. Stensby, C. R. Christensen, J. A. Grisham, and F. W. Clarke, U. S. Army Missile Command Technical Report RD-WS-91-1, August 1991.
3. R. K. Ahrenkiel, P. Weiss, D. Watkins, S. K. Gulati, and W. W. Grannemann, *J. Appl. Phys.* **49**, 2265 (1978).
4. D. Long and J. L. Schmit, *Semiconductors and Semimetals* **5**, 175, eds. Willardson and Beer, (Academic Press, 1970).
5. T. S. Moss, *Phys. Stat. Sol.* **2**, 601 (1962).
6. S. D. Smith, T. S. Moss, and K. W. Taylor, *J. Phys. Chem. Solids* **11**, 131 (1959).
7. G. A. Tanton, C. R. Christensen, J. A. Grisham, and J. A. Stensby, U. S. Army Missile Command Technical Report CR-RD-RE-90-1, March 1990.
8. J. A. Mroczkowski and D. A. Nelson, *J. Appl. Phys.* **54**, 2044 (1983).

*CHU

A More Rational and Rigorous Analysis of In-Bore Projectile

*Shih C. Chu, PhD

Simulation & Evaluation Br, Light Armament Div, CCAC
U.S. Army Armament Research, Development and Engineering Center
Picatinny Arsenal, New Jersey 07876-5000 (Ph 201-724-7316)

In contrast to the traditional designing and analyzing techniques (theories of elasticity), the plastic flow in the projectile has been taken into account. On the basis of Drucker's flow rule of plasticity, the Von Mises flow criterion, and the strain-hardening and compressibility properties of a material, a nonlinearity elastoplastic method has been developed for analyzing a projectile subjected to actual complicated firing conditions. Geometric nonlinearity has been considered by using the finite element technique. An incremental loading procedure has been used to consider the actual firing environment of a gun. The complete loading history which includes propellant gas pressure, rotating band pressure, setback force, and spin rate of a projectile was defined as a function of time.

INTRODUCTION

Projectiles are traditionally analyzed and designed by using theory of elasticity (1,2,3). In this approach, the entire projectile body is considered to be loaded within the elastic region. However, under actual firing conditions, the equivalent stress in some region of the projectile is much greater than the yield strength of the projectile material. Hence, plastic flow will be encountered in the projectile body. The present trend in stress analysis of weapon and ammunition components is to take into account, in more rigorous fashion, the complex phenomena of plastic flow. This is due to the necessity of designing for maximum stress to achieve an optimal design. In this investigation, a more rigorous nonlinear technique is developed in order to predict the inelastic deformation and stress distribution of a projectile subjected to actual firing conditions. Both nonlinear material response and geometric nonlinearity have been taken into consideration. Nonlinearity of material properties has been taken account by use of theories of plasticity. Geometric nonlinearity has been considered by use the finite element approach. A complete inelastic stress analysis of a 175mm, M437, HE projectile subjected to Zone 3 charge has been conducted. A load history was defined which subjected the projectile to the loads at various time on the travel curve.

CONSTITUTIVE EQUATIONS

The primary objective of this investigation is to develop a more rational, nonlinear elastoplastic method for analyzing and designing a projectile to achieve in-bore structural integrity. The incremental stress-strain relations associated with the Von Mises yield criterion obeying the kinematic-hardening law for work-hardening materials will be used. The kinematic hardening law in the incremental theory of plasticity was originally proposed by Prager [4], for the deformation of isothermal solids. Later, Prager [5] extended these formulations to the nonisothermal condition for rigid work-hardening solids. Recently, Chu [6,7], expanded Prager's work for solids of more general deformation state. A brief outline of the constitutive equations used in the analysis is now given.

The total increment strain tensor, $d\epsilon_{ij}$ is assumed to be the sum of elastic part, $d\epsilon_{ij}^e$, plastic part, $d\epsilon_{ij}^p$, and the part of thermal strain, $d\epsilon_{ij}^T$, i.e.

$$d\epsilon_{ij} = d\epsilon_{ij}^e + d\epsilon_{ij}^p + d\epsilon_{ij}^T \quad (1)$$

The elastic strain components $d\epsilon_{ij}^e$ are related to the incremental stress components, $d\sigma_{ij}$, by

$$d\sigma_{ij} = D_{ijkl} d\epsilon_{kl}^e = D_{ijkl} (d\epsilon_{kl} - d\epsilon_{kl}^p - \beta dT \delta_{kl}) \quad (2)$$

in which β is the thermal expansion coefficient, dT , is the change in temperature, and δ_{kl} is Kronecker delta. For isotropic material, the 4th rank material tensor, D_{ijkl} is defined by

$$D_{ijkl} = \frac{E}{1+\nu} \delta_{ik} \delta_{jk} + \frac{\nu E}{(1+\nu)(1-2\nu)} \delta_{ij} \delta_{kl} \quad (3)$$

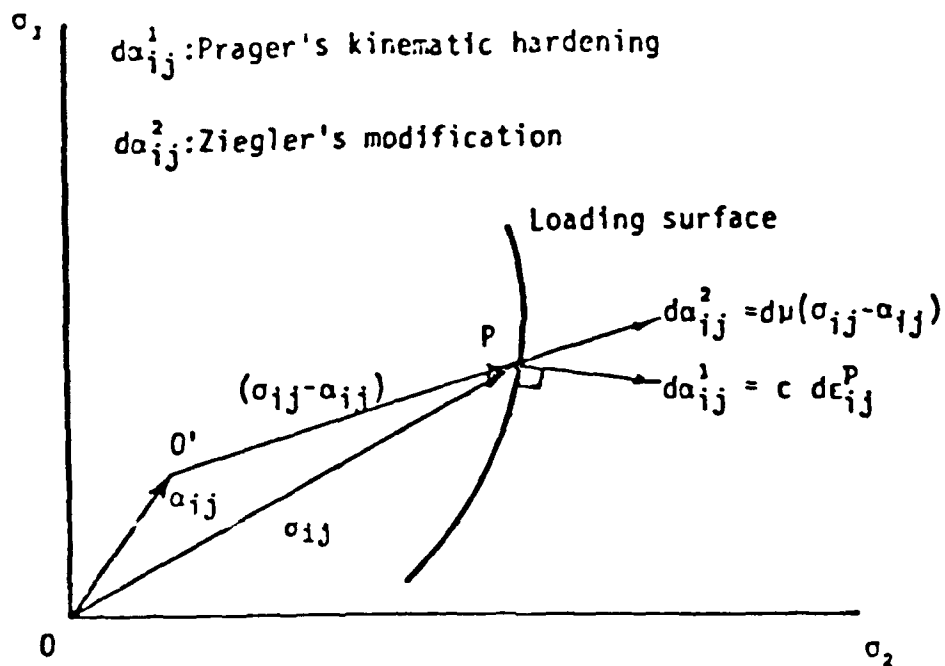
where E is Young's modulus and ν is Poisson's ratio of material.

On the basis of the Von Mises yield criterion with temperature-dependent yield strength of a material, the yield surface can be represented as:

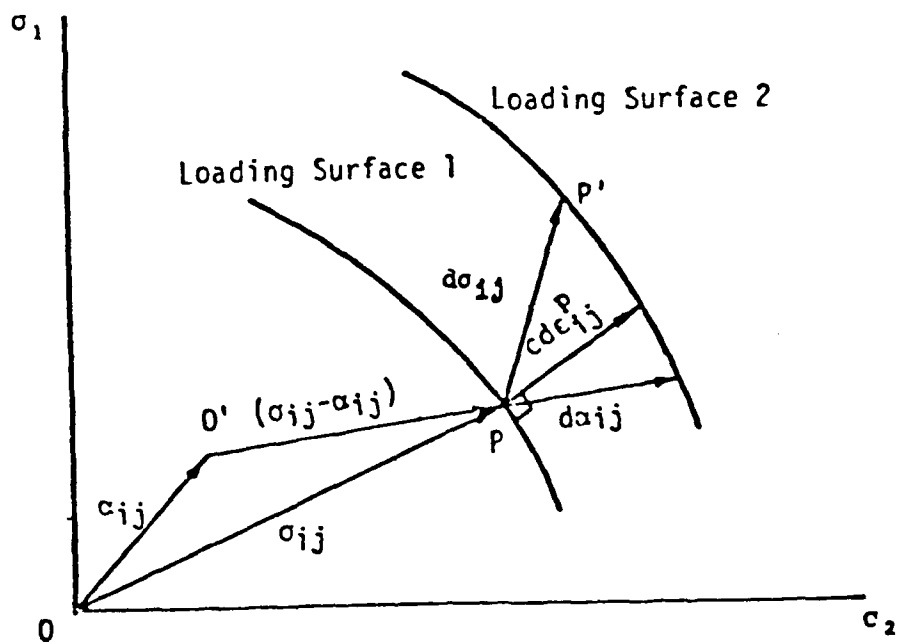
$$f = \frac{1}{2} S_{ij} S_{ij} - \kappa^2(T) = 0 \quad (4)$$

where

$$S_{ij} = (\sigma_{ij} - \alpha_{ij}) - \frac{1}{3} (\sigma_{kk} - \alpha_{kk}) \delta_{ij}; \quad S_{mm} = 0 \quad (5)$$



(a) Comparison of Prager's kinematic hardening rule with Ziegler's Modification



(b) Hardening rule and Definition of Plastic Strain Increment for strain-hardening material

Figure 1. Kinematic Hardening Rules

CHU

α_{ij} is a tensor representing the total translation of the center of the initial yield surface, and κ is related to the uniaxial yield stress $\kappa = \sigma_y(T)/\sqrt{3}$.

In addition to the yield condition, a constitutive relation between plastic strain increments, stress, and stress increments is required to describe the inelastic behavior of a material. The constitutive relation (flow rule) used in this investigation is based on Drucker's postulate for work-hardening material [8]. The flow rule is given as:

$$d\epsilon_{ij}^P = d\lambda \frac{\partial f}{\partial \sigma_{ij}} \quad (6)$$

where $d\lambda$ is a positive scalar quantity. On the basis of Drucker's statement, this plastic strain increment tensor must lie on the outward normal to the yield surface at the instantaneous stress state.

Based upon Prager's kinematic-hardening rule [4,9] with Ziegler's modification [10], the increment of translation of the center of the yield surface is assumed to be directed along the radius vector connecting the center of the yielding surface to the instantaneous stress state, i.e.,

$$d\alpha_{ij} = (\sigma_{ij} - \alpha_{ij}) d\mu; \quad d\mu > 0 \quad (7)$$

where $d\mu$ can be determined provided the stress point remains on the translated yield surface during plastic flow, i.e.,

$$d\mu = \frac{\frac{\partial f}{\partial \sigma_{kl}} d\sigma_{kl} + \frac{\partial f}{\partial T} dT}{(\sigma_{mn} - \alpha_{mn}) \frac{\partial f}{\partial \sigma_{mn}}} \quad (8)$$

A comparison of Prager's hardening rule and Ziegler's modification was shown in Figure 1 (a).

During plastic loading, the consistency condition requires that

$$df = \frac{\partial f}{\partial \sigma_{ij}} (d\sigma_{ij} - C d\epsilon_{ij}^P) + \frac{\partial f}{\partial T} dT = 0 \quad (9)$$

The plastic strain vector $C d\epsilon_{ij}^P$ is considered as the projection of $d\sigma_{ij}$ (and thus of $d\alpha_{ij}$) on the exterior normal to the yield surface, where C is a material constant. Hence, for small incremental of stress and strain, one can readily find that

CHU

$$d\lambda = \frac{1}{C} \frac{\frac{\partial f}{\partial \sigma_{ij}} d\sigma_{ij} + \frac{\partial f}{\partial T} dT}{\left(\frac{\partial f}{\partial \sigma_{mn}}\right) \left(\frac{\partial f}{\partial \sigma_{mn}}\right)} \quad (10)$$

Therefore, the flow rule becomes

$$d\epsilon_{ij}^P = \frac{\frac{\partial f}{\partial \sigma_{ij}}}{C \left(\frac{\partial f}{\partial \sigma_{mn}}\right) \left(\frac{\partial f}{\partial \sigma_{mn}}\right)} \left(\frac{\partial f}{\partial \sigma_{kl}} d\sigma_{kl} + \frac{\partial f}{\partial T} dT \right) \quad (11)$$

If axisymmetric deformation is considered, with reference to the cylindrical coordinates (r, θ, z) the state is defined by the nonvanishing stress components $\{d\sigma\}^T = \langle d\sigma_r, d\sigma_\theta, d\sigma_z, d\tau_{rz} \rangle$ and strain components $\{d\epsilon\}^T = \langle d\epsilon_r, d\epsilon_\theta, d\epsilon_z, d\gamma_{rz} \rangle$. Then, the incremental stress-strain relations are found in the following matrix form:

$$\{d\sigma\} = [\bar{D}] \{d\epsilon\} - \beta dT \{B\} \quad (12)$$

in which,

$$[\bar{D}] = \frac{\lambda}{\nu} \begin{bmatrix} 1-\nu & \nu & \nu & 0 \\ \nu & 1-\nu & \nu & 0 \\ \nu & \nu & 1-\nu & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix} - \eta_1 \cdot \begin{bmatrix} S_r^2 & S_r S_\theta & S_r S_z & S_r \tau_{rz} \\ S_\theta S_r & S_\theta^2 & S_\theta S_z & S_\theta \tau_{rz} \\ S_z S_r & S_z S_\theta & S_z^2 & S_z \tau_{rz} \\ \tau_{rz} S_r & \tau_{rz} S_\theta & \tau_{rz} S_z & \tau_{rz}^2 \end{bmatrix} \quad (13)$$

$$\{B\} = \frac{E}{1-2\nu} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{\eta_2}{\beta} \cdot \frac{\partial \kappa}{\partial T} \begin{bmatrix} S_r \\ S_\theta \\ S_z \\ \tau_{rz} \end{bmatrix} \quad (14)$$

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad (15)$$

$$\eta_1 = \frac{1}{g} \left(\frac{E}{1+\nu} \right)^2 \quad (16)$$

$$\eta_2 = \frac{2\kappa}{g} \cdot \frac{E}{1+\nu} \quad \text{and,} \quad (17)$$

$$g = 2\kappa^2 \cdot \left(\frac{E}{1+\nu} + C \right) \quad (18)$$

PROJECTILE CONFIGURATION AND MATERIAL PROPERTIES

The outline of a 175mm, HE, M437 projectile is shown in Figure 2. The projectile is made of steel. The material properties and stress-strain relation considered are shown in Figure 3.

LOADING CONDITIONS

The loads considered in this investigation simulate the environment in the gun barrel during firing. The four types of loading considered are shown in Figure 4:

- 1) Propellant gas pressure
- 2) Rotating band pressure
- 3) Setback force due to acceleration of projectile
- 4) Rotational velocity due to the spin of projectile

The frictional shear forces between the band and the barrel were neglected in this investigation. The loads were applied in incremental fashion; each load point simulated the physical interdependence of the loads at different times during the interior ballistic cycle. The detailed loading history for a 175mm HE, M437 projectile subjected to Zone 3 charge is shown in Figure 5.

METHOD OF SOLUTION

The projectile and all loading acting on the projectile are considered as axisymmetric. A cylindrical coordinate system (r, θ, z) is used in this analysis. By the assumption of axial symmetry, all variables are independent of angle θ ; consequently, all derivatives, with respect to θ vanish. The displacement u_θ , and the shear stresses $\tau_{r\theta}$ and $\tau_{\theta z}$ vanish.

Due to the complexity of the geometry of a projectile and nonlinear material behavior, the finite element method (11, 12) was used to conduct the stress analysis. The analytical approach used in this investigation is the incremental loading technique, wherein at each step of loading a new stiffness matrix is formulated, in terms of the finite element model, and solved for incremental deformations, stresses, and strains.

To perform a finite element stress analysis, the cross section of a projectile is divided into a large number of small triangular and

CHU

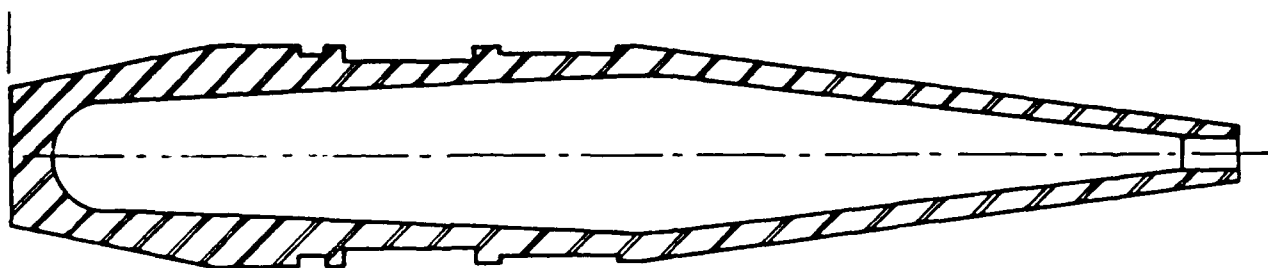


Figure 2. Projectile Configuration

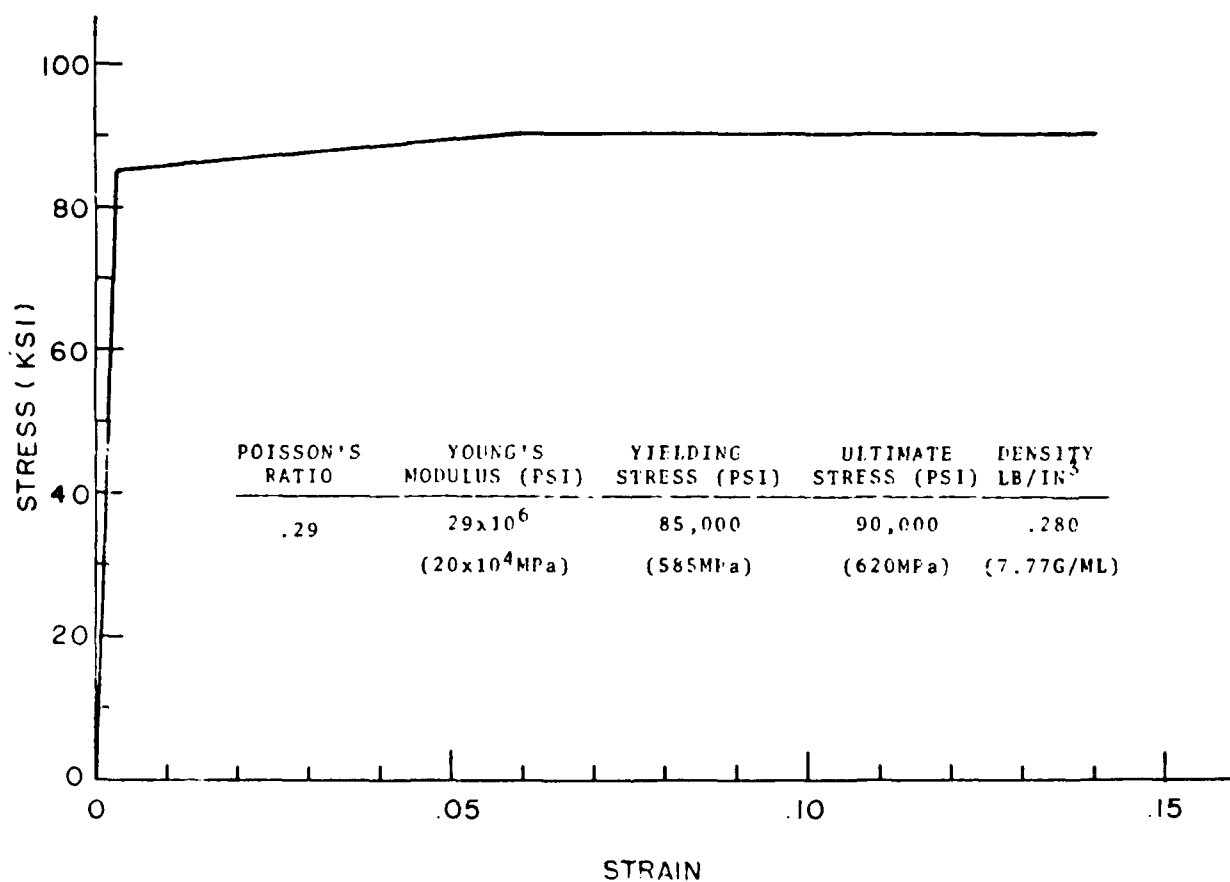


Figure 3. Stress-Strain Relation and Material Properties for Projectile Material

CHU

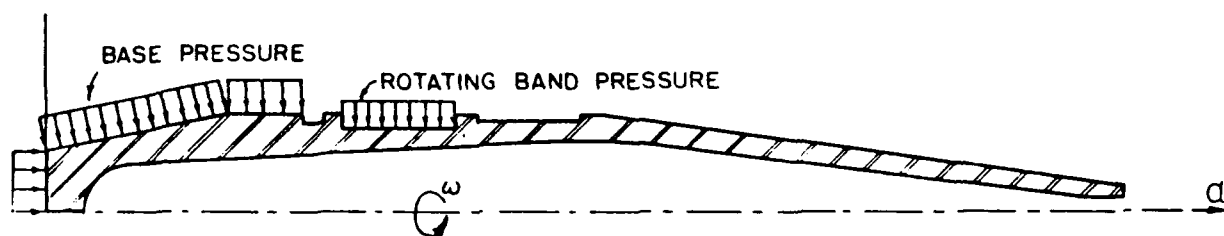


Figure 4. Loading Conditions

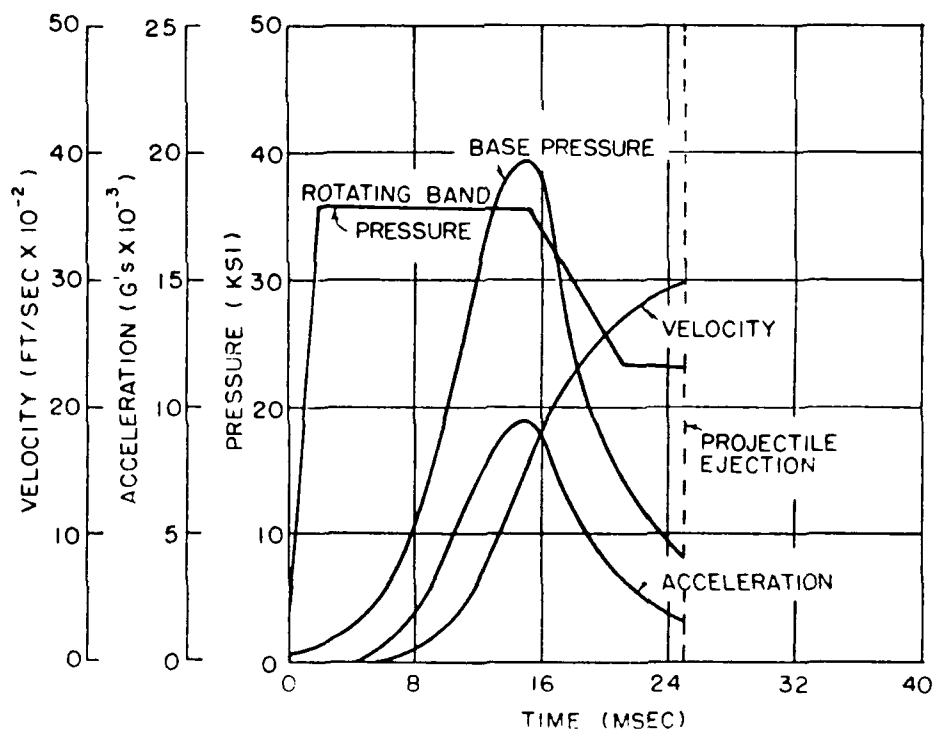


Figure 5. Loading History - A 175 mm, HE M437 Projectile Subjected to Zone 3 Charge

CHU

quadrilateral elements as shown in Figure 6. Only a few nodal points on the boundary need be specified; the remaining nodal points are obtained from an automatic mesh generating computer program. The grid was partitioned to place a finer grid at those areas of the body that are expected to undergo large stress.

COMPUTATIONAL RESULTS

The computational results presented in this investigation are based upon the following three major loading conditions as shown in Figure 5:

- 1) When rotating band pressure reaches the maximum value of 36,000 psi (245 MPa).
- 2) When both propellant gas pressure and setback force reach the peak.
- 3) At the time the projectile is ejected when the spin rate reaches the maximum at the muzzle end of the tube.

For each incremental of loading, the stresses and strains of each element and the displacement of each nodal point were calculated. The plastic flow has initiated within elements directly under the rotating band in the inner surface of the projectile at rotating band pressure reaching 25,000 psi (172 MPa). Based upon the Von Mises yield criteria, the equivalent stress for each element was calculated.

At maximum rotating band pressure, the material in the entire range below the rotating band is yielded. The plastic range in the projectile is indicated in Figure 7. The deformed configuration under maximum rotating band pressure is shown in Figure 8 by solid lines. The equivalent stress distribution along the projectile length at inner and outer surfaces is shown in Figure 9. The maximum equivalent stress has reached 85,300 psi (578 MPa) which is a little higher than the yield strength ($\sigma_y = 85,000$ psi = 584 MPa) and lower than the ultimate strength ($\sigma_u = 90,000$ psi = 620 MPa) of projectile material. The maximum plastic strain encountered is 0.004 in/in, which is relatively small.

At maximum propellant gas pressure and maximum setback force, the plastic flow has taken place in several places in the projectile as shown by the shaded areas in Figure 10. The deformed configuration is shown in Figure 11 by solid lines. The dashed line configuration is the undeformed configuration. The equivalent stress distribution along the projectile length at inner and outer surfaces is shown in Figure 12. The plastic range has been expanded several places between $z = 1.6$ inches (4.06 cm) and $z = 17.3$ inches (44 cm). The maximum equivalent stress in the projectile is 85,300 psi (587 MPa). The maximum plastic strain is 0.004 in/in.

At the time the projectile reaches the muzzle end while the spin rate has attained the peak, the maximum equivalent stress has only reached

CHU

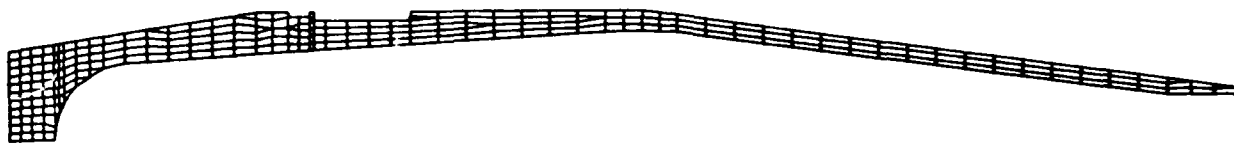


Figure 6. Finite Element Presentation of a Projectile



Figure 7. Plastic Zone in the Projectile at Maximum Rotating Band Pressure

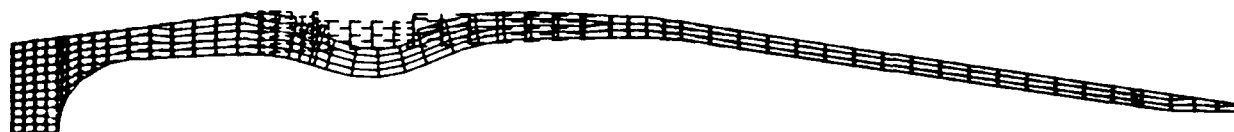


Figure 8. Deformed Grid Mesh at Maximum Rotating Band Pressure

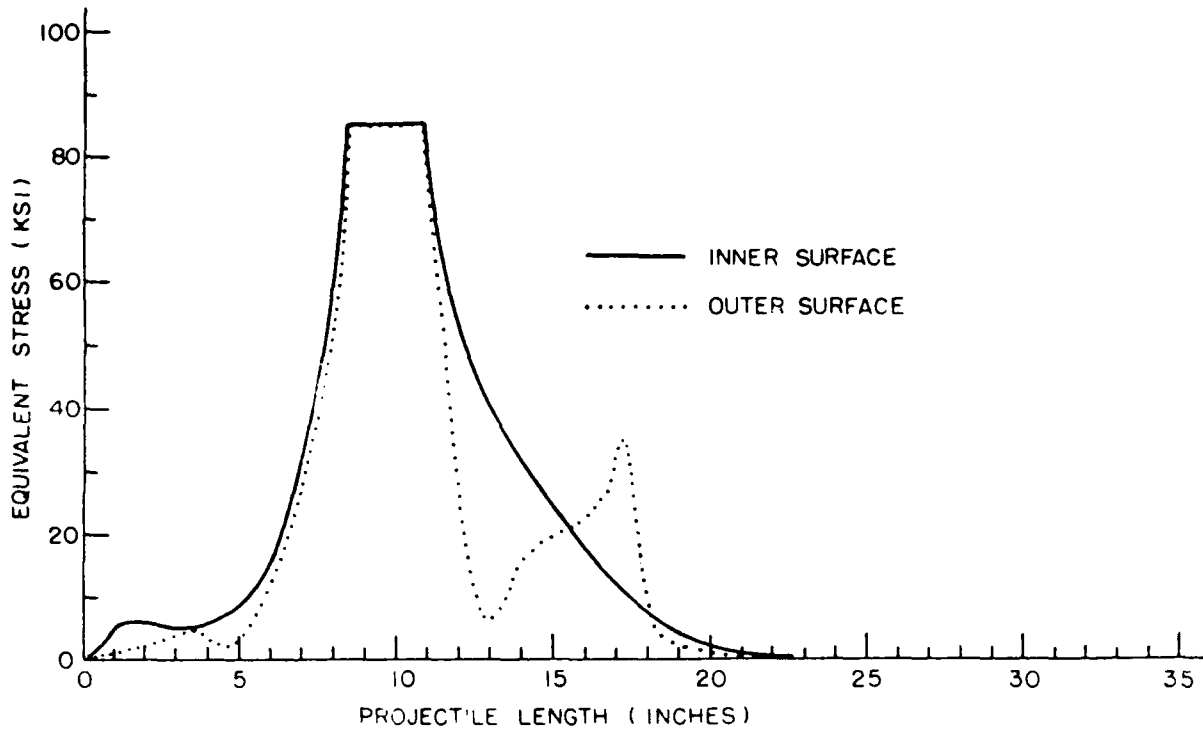


Figure 9. Equivalent Stress Distribution Along the Projectile While the Rotating Band Pressure Reaches the Maximum



Figure 10. Plastic Zone in the Projectile at Maximum Propellant Gas Pressure and Maximum Setback Force

CHU

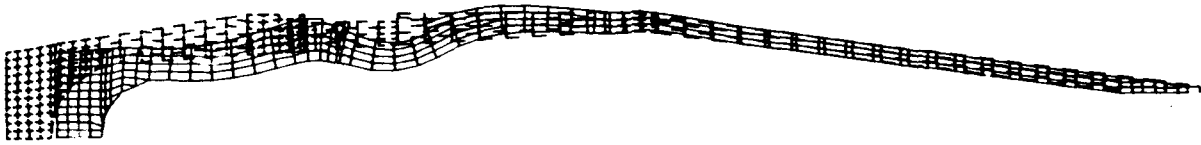


Figure 11. Deformed Grid Mesh at Maximum Propellant Gas Pressure and Maximum Setback Force

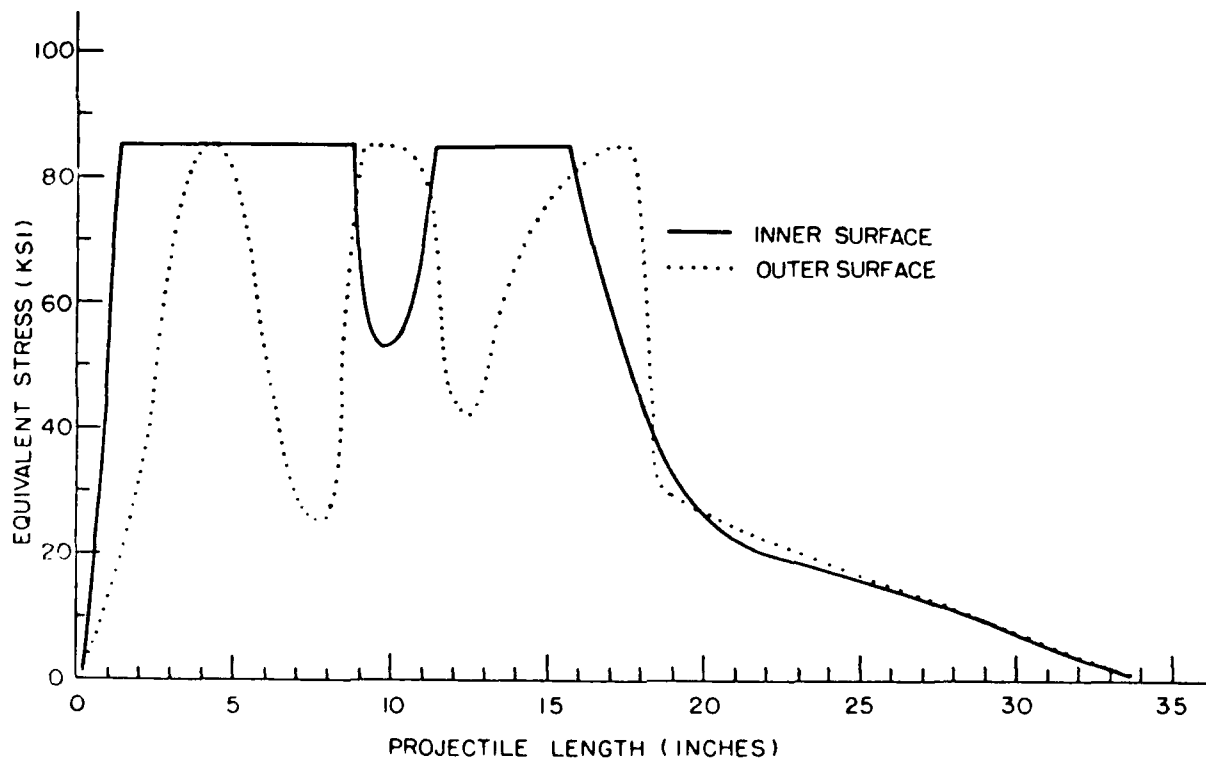


Figure 12. Equivalent Stress Distribution Along the Projectile at Maximum Propellant Gas Pressure and Maximum Setback Force

CHU

46,000 psi; hence the entire projectile body is in elastic range. The deformed configuration is shown in Figure 13.

The maximum equivalent stress of 85,300 psi (587 MPa) was reached during the entire loading process. This value is greater than the yielding strength ($\sigma_y = 85,000$ psi = 584 MPa) and smaller than the ultimate strength ($\sigma_u = 90,000$ psi = 620 MPa) of the projectile material. If the investigation was based on elastic analysis, the maximum equivalent stress could reach a value which would be much higher than the ultimate strength of the projectile material. This is the reason that a nonlinear analysis has been adopted, since a more rigorous stress analysis can only be obtained by using nonlinear theories of plasticity.

CONCLUSIONS

On the basis of Drucker's flow rule of plasticity, the Von Mises flow criterion, and the strain-hardening and compressibility properties of a material, a more rational rigorous nonlinear elastoplastic analytic method has been developed for analyzing and designing a projectile subjected to actual complicated firing conditions. In contrast to the traditional designing and analyzing techniques (theories of elasticity), the complex phenomenal plastic flow in the projectile has been taken into account.

Both nonlinear material response and geometric nonlinearity have been taken into consideration. Nonlinearity of material properties has been taken into account by use of theories of plasticity. Geometric nonlinearity has been considered by use of the finite element technique. An incremental loading procedure has been used to simulate the actual firing environment of a gun. The complete loading history which includes propellant gas pressure, rotating band pressure, setback force, and spin rate of a projectile was defined as function of time.

At each incremental of loading, the critical region is determined. In this region, the equivalent stress in general is above the yield strength of material; however, it is below ultimate strength of the material. The plastic strain has taken place in the region; however, maximum equivalent plastic-strain at inner surface is below 0.004 in/in, which is relatively small.

Based upon this investigation, the plastic zone in the projectile has been identified and accurately located for each incremental of loading which simulates the actual fire environment in a gun. The developed technique will provide a more rigorous stress analysis tool for projectile design. The potential design flaws and critical regions of a projectile can be identified before the projectile is made and tested.

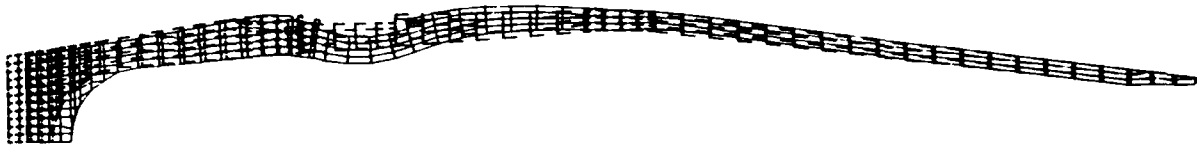


Figure 13. Deformed Grid Mesh as the Spin Rate Reaches the Maximum

REFERENCES

1. DePhillipo, T.E., and Booth, A.W., "Mathematical Model for Determining Stresses in Projectile Bodies", Report R-1939, Frankford Arsenal, Oct. 1969.
2. Mechanical Engineering Department, New York University, "Principal and Combined Stresses of the Shell, HE 155mm M197", Picatinny Arsenal Report, U38115.
3. Elder, A.S., Burns, B.P., and Hurban, J.M., "Stress Analysis of 175mm Projectile, HE M437", BRL Memorandum Report 2113.
4. Prager, W., "The Theory of Plasticity: A Survey of Recent Achievements (Hames Clayton Lecture)", Proc. Inst. Mech. Eng. 69 (1955).
5. Prager, W., "Nonisothermal Plastic Deformation", Proc. Konink. Ned. Akad., Van Weeten, Series B, 61, NO (3), 1958.
6. Chu, S.C., "Nonisothermal Elastoplastic Deformation of Work-Hardening Solids", Paper presented at Joint ASME/ASCE Mechanics Conference, University of Colorado, 22-24 June 1981.
7. Chu, S.C., "An Incremental Approach to Nonisothermal Elastic-Plastic Deformation", Proc. 27th Conference of Army Mathematicians, 1981.
8. Drucker, D.S., "A More Fundamental Approach to Plastic Stress-Strain Relations", Proc. 1st U.S. National Congress Applied Mechanics, New York 1952.
9. Prager, W., "A New Method of Analyzing Stresses and Strains in Work-Hardening Plastic Solids", Journal App. Mech. Vol. 23, 1956.
10. Ziegler, H., "A Modification of Prager's Hardening Rule", Quarterly of Appl. Math. Vol. 17, 1959.
11. Zienkiewicz, O.C., "The Finite Element Method in Engineering Science", McGraw-Hill, 1971.
12. Marcal, P.V., and King, I.P., "Elastic-Plastic Analysis of Two-Dimensional Stress Systems by the Finite Element Method", Inter. J. Mech., Vol. 9, No. 3, 1967.

COOKE, CHANG, TAYLOR, EVALDSSON

Strained Quantum Well Field Effect Lasers for Optoelectronic Integration.

P.W.Cooke*,

Geo Centers Operations at Ft. Monmouth,

Ft Monmouth, NJ 07703-5016

W. Chang,

US Army ETDL, SLCET_ED,

Ft. Monmouth, NJ 07703-5016,

G.W. Taylor, P.A. Evaldsson,

AT&T Bell Laboratories,

Holmdel, NJ 07733-3030

Considerable emphasis is currently being placed upon the ability of high speed optical links to replace existing electrical coaxial cables in systems such as phase array radars ¹. Such applications place considerable demands upon the devices employed to perform the optoelectronic conversion at both the transmitter and receiver ends of the link. In particular, for the transmitter elements, high speed and high efficiency semiconductor lasers would be desirable. Future systems have the potential to require large numbers of such devices. From both cost and complexity considerations, integration of lasers and either other optical elements or drive electronics would be advantageous for these applications.

More esoteric functions can be envisaged for integrated circuits containing many optical transmitter and receiver units. Optical preprocessing has been proposed as a method

to speed up and enhance many complicated digital processing tasks.² Examples of such systems include pattern recognition and target identification. In these cases, the input data is intrinsically parallel in nature, so that filtering or preprocessing with the input still in parallel form would be beneficial. With suitably adaptable circuits, arrays of "smart pixels," could be envisaged that use in-plane electronic processing and plane to plane multiple optical interconnects. Vertically stacking several planes would then provide the potential for a full parallel processing arrangement. In an ideal situation, even the output of the system could be in optical form.

Several different approaches have been tried to integrate lasers with either FET or bipolar devices. The most common is to employ a multiple or stacked growth technique³, where the laser structures are grown either above or below the transistor structures. This intrinsically creates problems of planarization and interconnection. An alternative is to use a common epitaxial layer sequence that enables fabrication of the different devices from the same basic structure. Appropriate choice of processing steps then ensures that both lasers and drive or logic electronics can be integrated at will.

Recently, the Heterojunction Field Effect Laser (HFEL) was demonstrated in the AlGaAs/GaAs material system⁴. The epitaxial layer sequence for the HFEL is totally compatible with that of the Heterojunction Field effect transistor (HFET)⁵. Both the HFET and HFEL are formed by placing a heavily doped charge sheet adjacent to an isotype heterojunction. With the correct choice of dopings, an inversion channel is formed at the interface. Modulation of the inversion charge is achieved by using an ohmic gate contact. HFEL operation is obtained by configuring the doping profile to produce the inversion charge in the multiquantum well (MQW) active region of a suitable optical waveguiding structure. Due to the use of the ohmic contact, the gate contact of the HFET can be spatially removed

from the channel to enable a waveguiding region to be incorporated. For the N-channel configuration, biasing the gate-source diode beyond breakdown provides hole injection from the gate contact and electron injection into the channel for lasing. As such, the HFEL effectively functions as a lateral current injection laser. The intrinsic compatibility of both devices ensures that they can both be fabricated alongside each other with the appropriate choice of processing steps. Thus they naturally lend themselves to the prospect of optoelectronic integration.

In this paper, results of fabrication of the HFEL and HFET using the strained InGaAs/AlGaAs material system are presented. Compared to the lattice matched GaAs/AlGaAs material system, incorporation of coherently strained InGaAs quantum wells offers several advantages. The incorporated strain significantly modifies the valance band structure of the Quantum well and reduces the in-plane hole mass. The lifting of the zone center degeneracy reduces the density of states function and has been demonstrated to produce lower threshold current densities and faster small signal response in conventional diode lasers⁶. The smaller bandgap also means the GaAs substrate is transparent at the emission wavelengths, an important consideration for the optical preprocessing applications envisaged above. From the HFET perspective, the use of strained MQW structures has the potential to increase the maximum channel charge that can be obtained prior to parasitic parallel conduction occurring in the immediate gate ("barrier") region.

The basic growth structure used for these experiments is shown in Fig 1. The entire structure was grown using Molecular Beam Epitaxy on a semi insulating GaAs substrate. The growth temperature for the AlGaAs and GaAs layers was 580 C and for the InGaAs quantum well regions , 525 C. The growth temperatures for the AlGaAs layers are much lower than those conventionally considered necessary for optimum laser performance⁷. The use of the

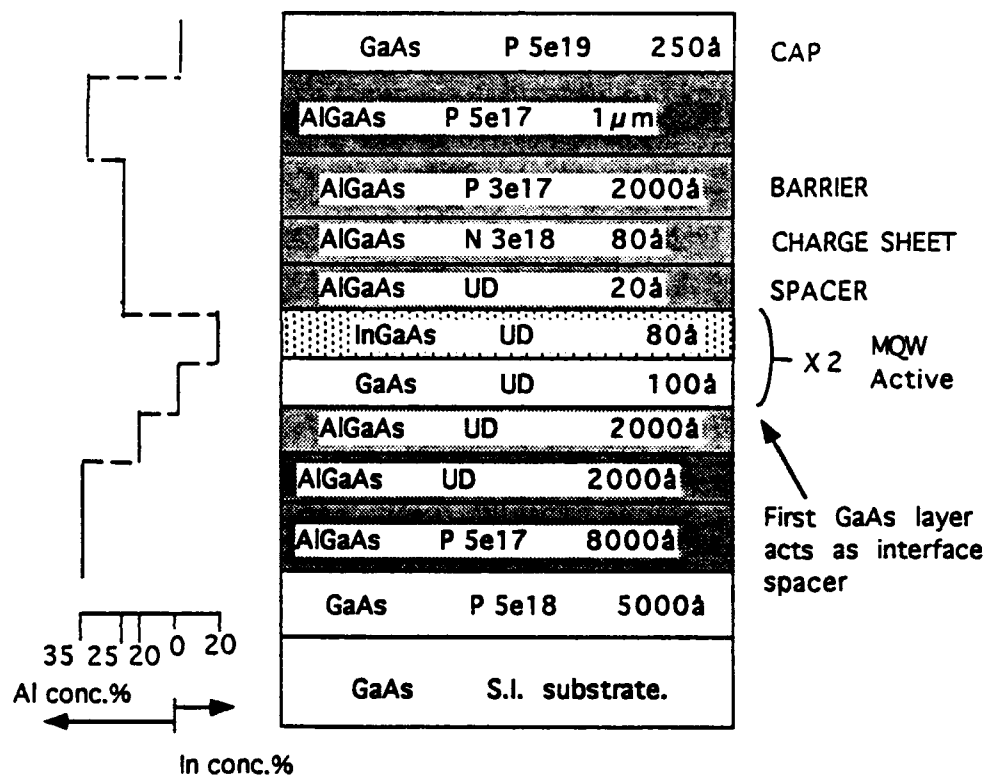


Fig.1. Epitaxial layer sequence for the HFEL/HFET

strained MQW region with GaAs barriers, means the usually poor lower AlGaAs/GaAs interface can be spatially separated from the MQW active region. Thus non-radiative parasitic interface recombination can be limited. Crucial to maintaining high luminescence efficiency is the correct choice the binary (GaAs) spacer thickness. Typically a 100-200Å GaAs spacer has been found necessary. To facilitate the change in growth temperatures for the InGaAs layers, a 2 min growth interruption is employed. The use of lower overall growth temperatures also has benefits in terms of dopant stability, since the HFET voltage threshold is determined exclusively from the dopant profile within the structure (unlike for example conventional modulation doped FET [MODFET] devices which rely on gate recess procedures). Control over the FET parameters, such as the voltage threshold, becomes crucial when larger scale logic circuits are considered

The detailed structure consists of a heavily doped charge sheet (N+) grown above the strained InGaAs MQW region. A small AlGaAs spacer region is incorporated in a similar manner to MODFET structures. The charge sheet is depleted both at equilibrium and under bias. The region above the charge sheet is p-doped and is designated the barrier region. HFET operation can be viewed as the trade-off of barrier depletion charge and inversion channel charge. The surrounding layers form an optical waveguide in a similar manner to a conventional diode laser. The lower waveguide layer also acts as a collector or backgate contact for the HFET/HFEL. Key to operation of both the HFEL and HFET is the use of a highly P+ doped cap layer as the gate or emitter contact.

A schematic device cross section is shown in Fig.2. Fabrication of the HFEL proceeds by initially defining a refractory gate (stripe) contact. Lateral waveguiding is obtained by dry etching down to within 0.2µm of the active region. Using the refractory contact as an implant mask , Si ion implantation is used to form source and drain contacts.

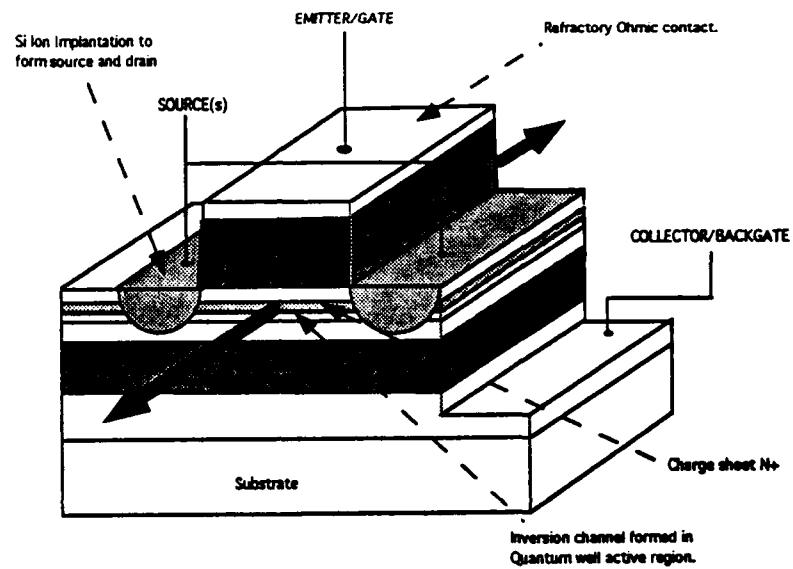


Fig.2. Schematic cross section HFEL/HFET.

Depending upon the exact configuration desired, the HFEL can single or dual side sourced. The devices tested here were single sourced for simplicity. Activation of the implants is achieved using a Rapid Thermal Anneal(RTA) at 950 C for 10 sec. The backgate or collector contact is formed by etching down below the MQW region. Final bond pad metalisation steps using polyimide isolation layers are used to enable contacting the device. Thinning and cleaving is performed in an identical manner to diode lasers to produce facets. HFET fabrication is similar, except for the gate (stripe) dimensions and final thinning step.

Operation of the HFEL is achieved by biasing the emitter contact with respect to the source contact. Typical pulsed (200nsec,1kHz rep.) light output characteristics are shown in Fig.3. Threshold currents of 34mA, 32mA and 22mA were obtained on 15, 10, and 5 μ m wide devices. The corresponding threshold current densities were 500,710 and 970A/cm² respectively. Typical slope efficiencies were 0.25W/A/facet. Continuous operation of these devices was obtained with threshold currents of 40mA,38mA and 29mA. Maximum C.W. output was limited due to the device mounting arrangement and higher than expected series resistance(15 Ω). This in turn was attributed to a higher than expected source resistance term resulting from non-optimization of the implant conditions. With many lateral current injection lasers, there is a possibility of filamentary action from non-uniform current injection across the stripe. To confirm whether this was happening, the far field pattern of a 10 μ m wide HFEL was recorded. This is shown in Fig.4. From the observed symmetry, *uniform current injection across the stripe can be inferred*. The wider 15 μ m HFEL's exhibited multimode operation as could be expected from the dimensions of the lateral waveguide structure.

The wavelength of emission of the HFEL was measured using an Anritsu spectrum analyzer. For C.W.operation, emission wavelengths were 1.02 μ m for 10 μ m wide stripes.

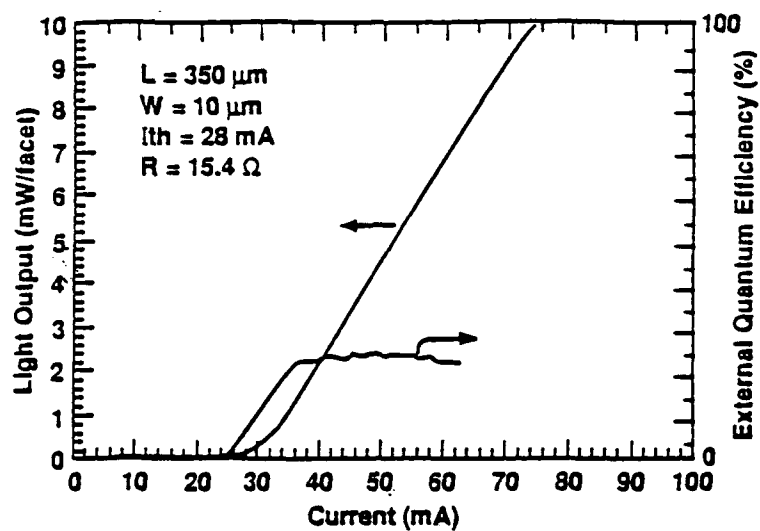


Fig. 3 Pulsed light output characteristics.

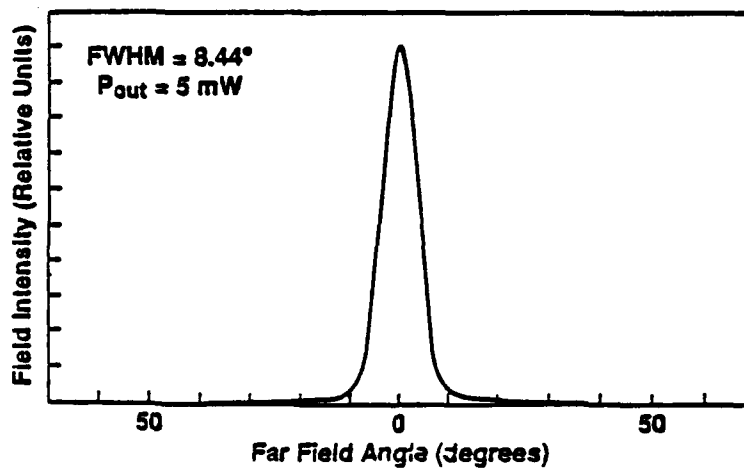


Fig.4. Measured Farfield pattern, parallel to junction.

The output spectra is shown in Fig5.

Small signal response of the HFEL was measured using a high speed InGaAs photodiode and a network analyzer. Due to the poor response of this particular photodiode, the measured signal to noise was not particularly good. The expected increase in cut-off frequency with bias level was observed and is shown in Fig.6. Clear relaxation oscillations can be seen in the output. The maximum-3dB cutoff frequency obtained was 4GHz at 6mW output (approximately $2xI_{th}$). Further increases in modulation bandwidth were limited due the falling C.W. output as a result of ohmic heating.

The measured DC drain output curves for the HFET fabricated from this material are shown in Fig.7a. Fig.7b shows the measured transconductance as a function of gate bias. Peak normalized transconductance is 87ms/mm for a nominal 2 μ m long, 50 μ m wide gate. The HFET threshold voltage was -0.3V. The maximum drain output current densities obtained were 120mA/mm.

The D.C. compatibility of the two devices can be determined using the HFET drain output and HFEL threshold currents. Simple scaling of the present results suggests the capability to run the smaller 5 μ m HFEL's to just over threshold assuming equal stripe length and gate width devices. Both the HFET and HFEL have considerable scope for optimization though. The HFET drain output can be increased by increasing the charge sheet doping and barrier region doping to ensure a greater maximum inversion charge. Both values necessarily have to change to ensure a accurate threshold voltage is maintained for the HFET. The limit to this process is determined by two factors. Firstly practical considerations regarding achieving high Si doping levels using MBE; and secondly, the available conduction band offset obtained at the quantum well/barrier interface. This directly determines the available number of states that can be filled prior to thermionic

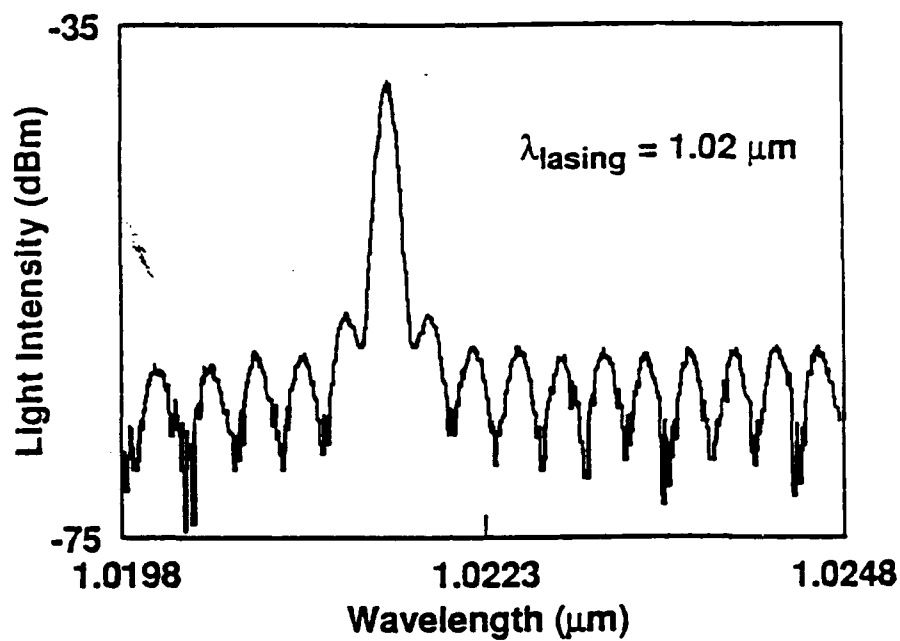


Fig.5. C.W. Output spectrum.

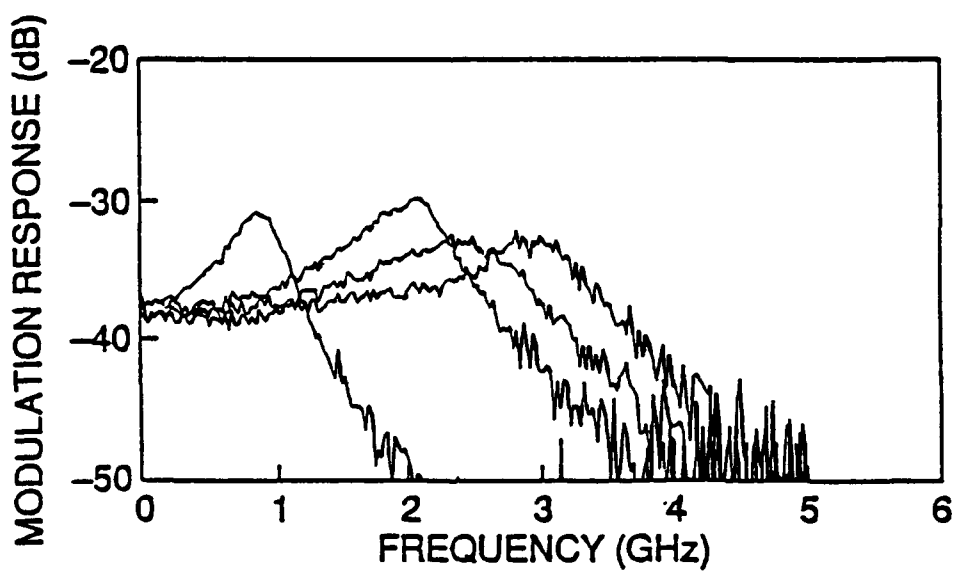
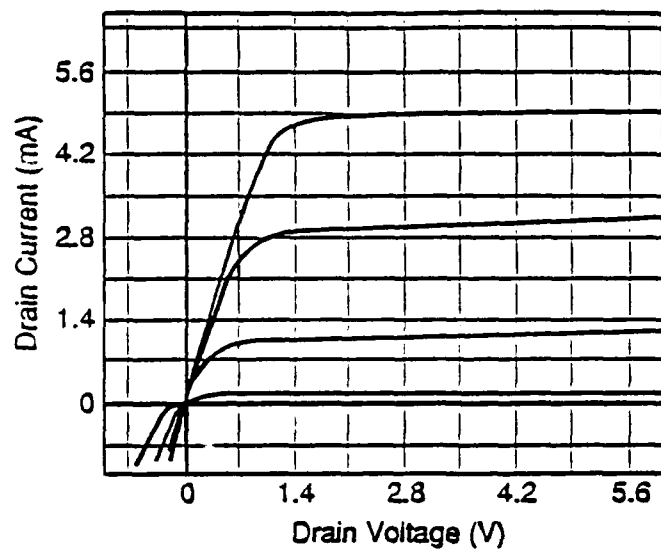
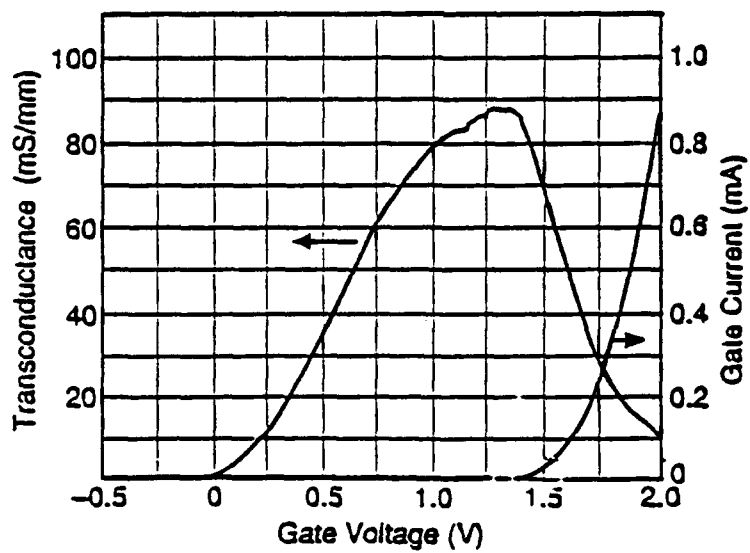


Fig.6. HFEL small signal response



(a)



(b)

Fig.7. a. HFET drain output characteristics.

b. HFET transconductance and gate current as a function of gate bias.

emission into the barrier becoming a problem. The strained InGaAs/AlGaAs material system is clearly advantageous in this respect when compared to the lattice matched GaAs/AlGaAs system. For comparison, strained MODFET, 550mA/mm channel currents have been reported for single channel structures⁸.

The need to make low resistance contact to the channel of the HFET means the etch depth and implant profile are not necessarily optimized for the HFEL lateral waveguiding. The effect of optical losses can clearly be seen from the increasing threshold current density with decreasing stripe width. Refinements to this part of the processing sequence should enable lower threshold currents to be realized. The HFEL output efficiency can be improved by modifying the waveguide structure to provide for tighter carrier confinement within the active MQW region. Higher MQW barriers and graded optical confinement regions could be incorporated relatively easily. The operation of the HFET does impose some constraints upon the surrounding AlGaAs layers though. Increasing the Al content of the barrier and collector layer immediately below the MQW active region would increase the confinement potential. However, due to the nature of Si donors in AlGaAs, Al mole fractions above 25% have a tendency to be resistive as a result of deep donor formation. These such layers would further compound the source resistance problem for both the HFEL and HFET. Improvements in the small signal modulation bandwidth will first require the series resistance, and in particular the source resistance to be lowered such that higher output powers can be achieved (comparable to the observed pulsed response). Thereafter reduction in threshold and improvements in efficiency should significantly enhance the small signal response, in an analogous manner to conventional diode lasers⁹.

In summary, an FET and lateral current injection laser have been grown and fabricated using a single epitaxial layer sequence. Initial results have established the

COOKE, CHANG, TAYLOR, EVALDSSON

concurrent operation of both devices using identical processing steps. Future integration should be possible to fulfill the potential of both multiple optical microwave links and smart pixel applications for optical preprocessing.

Acknowledgements.

The authors would like to thank B. Tell and K. F Brown Goebler for assistance with the ion implantation steps: C. A Burrus for the device bonding and P.A. Kiely and M.G. Young for assistance with the device testing.

References.

1. K.Y Lau. *Proceedings of Device Research Symposium*, Charlotte NC p365. 1991
2. S.R. Forrest Paper ThH2, *Tech. Digest, OSA annual meeting 1991*, San Jose, Ca.
3. S.D. Offsey et al. *Elec.Lett* **26**, p350 1990.
4. G.W.Taylor et al *Appl.Phys.Lett.* **58**, p666 1991
5. G.W. Taylor et al. *IEEE Elec.Dev.Lett.* **10** p88 1989
6. R. Nagarakan et al *Appl. Phys. Lett.* **58**(21) p2326 1991.
7. W.T.Tsang "Semiconductor lasers and photodetectors." in *"Physics and Technoogy of Molecular Beam Epitaxy."* Ed. E. H. C. Parker Plenum 1985.
8. L.D. Nguyen et al. *IEEE Trans. Elec. Dev.* **36** (5) p833 1989.
9. G.P. Agrawal and N.K. Dutta *"Long Wavelength Semiconductor lasers."* Van Nostrand 1986.

Transient Energy Transfer in Wide Gap Semiconductors

Steven S. Cotariu* and John E. LaSala

Photonics Research Center
U.S. Military Academy
West Point, NY 10996

INTRODUCTION

Wide gap semiconductors are normally transparent to visible light. When exposed to high intensity light, however, nonlinear effects such as two photon absorption (TPA) may occur. TPA, in turn, leads to a host of interesting secondary effects involving the free carriers generated. One of these effects is Transient Energy Transfer (TET), which can occur when two coherent beams, in a pump-probe experimental configuration, are focused in such a material. In the present work we utilize polycrystalline samples of ZnSe and ZnS, both II-VI semiconductors. At intensity levels in excess of 1 GW/cm^2 we find a significant amount of energy is scattered from the 532 nm pump pulse ($\sim 300 \mu\text{J}$, 25 ps) into the probe pulse ($\sim 100 \text{ nJ}$, 25 ps) in ZnSe. The probe gain is a function of both pump intensity and relative arrival times of pump and probe.

The TET theory is implemented on a personal computer. The numerical results are compared with experimentally obtained data to determine the viability of the model.

THEORY

Energy transfer from a strong pump to a weaker probe is modeled using the TET theory,¹⁻³ in which the free carrier generation mechanism is TPA. Two coherent pulses of wavelength λ crossing with an angle of 2θ develop a transverse intensity pattern whose spatial period Λ (Figure 1) is given by:⁴

$$\Lambda = \frac{\lambda}{2 \sin \theta}. \quad (1)$$

Inside the semiconductor, this intensity interference pattern generates a matching pattern of free electrons and holes through the TPA process. This spatially varying free carrier population produces a transverse index of refraction (phase) and absorption (amplitude) grating through the Drude-Lorentz model⁵ that extends into the medium. The phase effects dominate absorptive effects in ZnSe.

The TET process may be understood qualitatively as a two step process. The free carrier diffraction grating is generated by interference between the leading edges of the pulses, which then scatters the trailing edges into each other. The geometry of the situation automatically insures that the first diffraction order from each pulse is scattered into the other pulse. This scattered light suffers a 90 degree phase shift with respect to the zeroth order². As the two beams propagate through the medium, they accumulate phase-shifted light that changes their net phase (but not necessarily their amplitudes).

If the input beams are of unequal intensity, a net phase difference between the pump and probe beams develops with propagation distance in the crystal. The interference pattern shifts transversely in phase (toward the weaker beam) as the beams move deeper into the crystal. If the recombination of the free carriers in the semiconductor is instantaneous, the phase grating will keep pace with the shifting interference pattern and no net energy transfer occurs. However, if there is a finite recombination time, the phase lag between the interference pattern and the phase grating results in energy transfer from the strong to weak beam. Maximum energy transfer occurs for a phase difference of 90 degrees⁶.

The evolution of similarly polarized beams crossing in ZnSe is modeled by the following first order differential equations. These are derived by application of the slowly varying envelope approximation to the wave equation with sources that include third order nonlinearities in the fields (two photon absorption, nonlinear refraction) and interaction with the free carriers:⁷

$$\begin{aligned}
 \cos \theta \frac{dA^2}{dz} &= -2\gamma_1 N_2 AB \sin(\phi - \nu) - \left[\alpha + \beta_2 (A^2 + 2B^2) + \sigma_{eh} N_0 \right] A^2 \\
 \cos \theta \frac{dB^2}{dz} &= +2\gamma_1 N_2 AB \sin(\phi - \nu) - \left[\alpha + \beta_2 (B^2 + 2A^2) + \sigma_{eh} N_0 \right] B^2 \\
 \cos \theta \frac{d\phi}{dz} &= \gamma_1 N_2 \left(\frac{A}{B} - \frac{B}{A} \right) \cos(\phi - \nu) - \beta_1 (A^2 - B^2).
 \end{aligned} \tag{2}$$

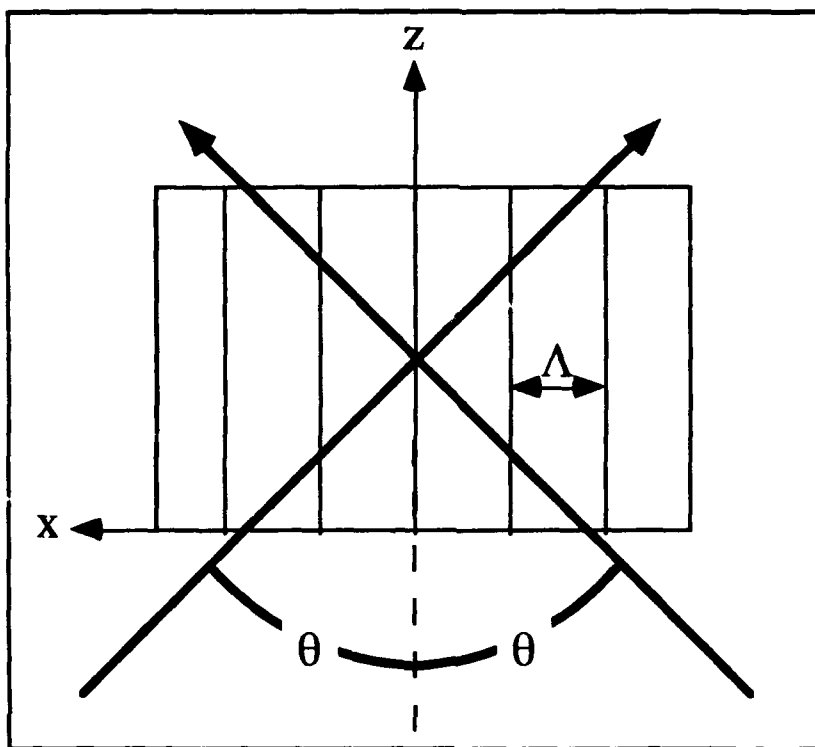


FIGURE 1

Fig.1: Interference pattern of two coherent beams of light crossing in a medium at an angle 2θ . The period of the intensity pattern Λ is related to the crossing angle and wavelength of light according to Equation 1. A matching pattern of free carriers will be generated through two photon absorption.

A and B are the amplitudes of pump and probe beams, respectively, $\phi = \phi_B - \phi_A$ is the difference in phase between probe and pump, and z measures the depth in the crystal along the bisector of the two beams. N_0 and N_2 are the Fourier components of free carrier densities as described in the next paragraph. The constants found in these equations and others below are given in Table 1.

The free carrier concentration is most conveniently described in terms of a "DC" background level N_0 and a modulation N_2 of period Λ and phase ν :

$$N = N_0 + 2N_2 \cos\left(2\pi \frac{x}{\Lambda} + \nu\right). \quad (3)$$

The x coordinate measures transverse position in the crystal. The free carrier concentration will evolve according to:⁷

$$\begin{aligned} \frac{\partial N_0}{\partial t} &= \frac{\beta_2}{2\hbar\omega} \left[(A^2 + B^2)^2 + 2A^2B^2 \right] - K_2(N_0^2 + 2N_2^2) \\ \frac{\partial N_2}{\partial t} &= \frac{\beta_2}{\hbar\omega} (A^2 + B^2) AB \cos(\phi - \nu) - 2K_2N_0N_2 \\ \frac{\partial \nu}{\partial t} &= \frac{\beta_2}{N_2\hbar\omega} (A^2 + B^2) AB \sin(\phi - \nu) \end{aligned} \quad (4)$$

Note that the quadratic loss terms included in the free carrier evolution equations assume the dominant form of recombination is radiative. A much slower linear recombination term has been ignored because we use short pulses. Also ignored is transverse carrier diffusion which would lower N_2 and raise N_0 . Over the time scale of the interaction transverse diffusion will not contribute significantly because of the low mobility of free carriers in ZnSe. Even less important is longitudinal diffusion. A future version of the model will accommodate this loss mechanism, however.

The pulses of light were taken to be Gaussian in time:

$$\begin{aligned} A &= A_0 \exp\left[-4 \ln 2 \frac{t^2}{\tau_p^2}\right], \\ B &= B_0 \exp\left[-4 \ln 2 \frac{(t - \Delta t)^2}{\tau_p^2}\right], \end{aligned}$$

where Δt is the time delay between pump and probe, and τ_p is the pulse length (FWHM).

The system of six coupled differential equations given by (2) and (4) are solved self-consistently using a fourth-order Runge-Kutta scheme. The pulses are sliced into time steps and numerically propagated through the 4.0 mm depth of the ZnSe

sample used. Stabilization of probe gain is used to signal convergence of the code as time and space steps are made smaller. The results of experimental beam characterization are applied to determine peak intensities of pump and probe and pulse length, and the values of the material constants from Table 1 are utilized. The parameters β_1 and K_2 are not well known for this material, so they are treated as adjustable parameters to bring the magnitude of the probe gain at optimal delay into agreement with experimental results. Adjustments are made from reasonable estimates of these quantities.

Cross polarized beams couple only weakly through direct TPA. Some coupling is possible because of the tensor nature of the third order susceptibility responsible for TPA, but the component connecting orthogonal polarizations is smaller than that for parallel polarizations in these materials. Therefore, except possibly at very high intensities, no gain is expected. Instead, the probe suffers increased linear absorption due to the large "DC" background of free carriers created by TPA of pump photons alone.

Ref.	Material Constant	Value
8	Index of refraction (n_0)	2.7
9	Attenuation coefficient (α)	0.6 cm ⁻¹
10	Nonlinear phase coeff. (β_1)	10 cm/GW
11	TPA coefficient (β_2)	5.5 cm/GW
12	"Drude" phase coeff. (γ_1)	1.6×10^{-16} cm ²
13	Free carrier cross-sec. (σ_{eh})	2×10^{-18} cm ²
14	Recombination rate (K_2)	5×10^{-8} cm ³ /s

TABLE 1

Table 1: Constants used in modeling Transient Energy Transfer in ZnSe. Uncertainties are implied by the least significant figure, except for K_2 which has a larger uncertainty (see reference).

EXPERIMENTAL CONFIGURATION

The experimental configuration is shown in Figure 2. The Continuum™ Nd:YAG laser output, at 10 Hz, is frequency doubled to 532 nm with an output energy of 35 mJ. Beam energy is controlled with variable attenuators and high power

beam splitters and measured using Laser Precision™ pulse energy detectors. A GBIP interface in a personal computer captures beam energies on a pulse-by-pulse basis. Probe arrival time is controlled with a retroreflector (corner cube) mounted on a translation stage with a position resolution of 10 μm (0.066 ps in arrival time).

The nominal temporal pulse length of the frequency-doubled beam is about 25 picoseconds based on estimates derived from the manufacturer's specifications. Pulse length is verified using two techniques. The first method employs a single-pulse autocorrelator (Continuum™) which encodes the temporal extent of a pulse along one spatial dimension. Using this instrument we measured the pulse length to be $\tau_p = (23 \pm 2)$ ps with negligible pulse to pulse variations. The second method employs the known nonlinear transmission characteristics of ZnSe to infer the pulse length. We assume the dominant non-linear attenuation mechanism of an intense pulse of light in ZnSe is due to two-photon absorption (TPA). The determination of the pulse length utilizing nonlinear transmission characteristics of ZnSe yields a value of $\tau_p = (24 \pm 2)$ ps. We employ a value of $\tau_p = 24$ ps for analysis.

Twenty measurements are made at each delay setting during the course of the experiment. For each shot the sampled energy (and through calibration, the pump and probe input energies), pump throughput energy, probe throughput energy, and delay stage setting are recorded. Probe throughput is recorded periodically with the pump blocked to establish a baseline for gain calculation.

RESULTS

The results of our gain study in ZnSe are shown in Figure 3. For each delay setting the twenty data sets were refined by rejecting input pulses that fell outside a window of about $\pm 5\%$ in energy, with a center energy of about 360 μJ for the input pump pulse. This translates to a peak pump intensity of about 3.6 GW/cm^2 . Typically half the pulses fell inside the window. The gain of each probe pulse fired with the pump present was calculated with respect to the probe throughput with pump blocked. The mean and standard deviation of the probe gain at each delay setting are plotted in Figure 3. Positive delay times correspond to arrival of the probe after the pump.

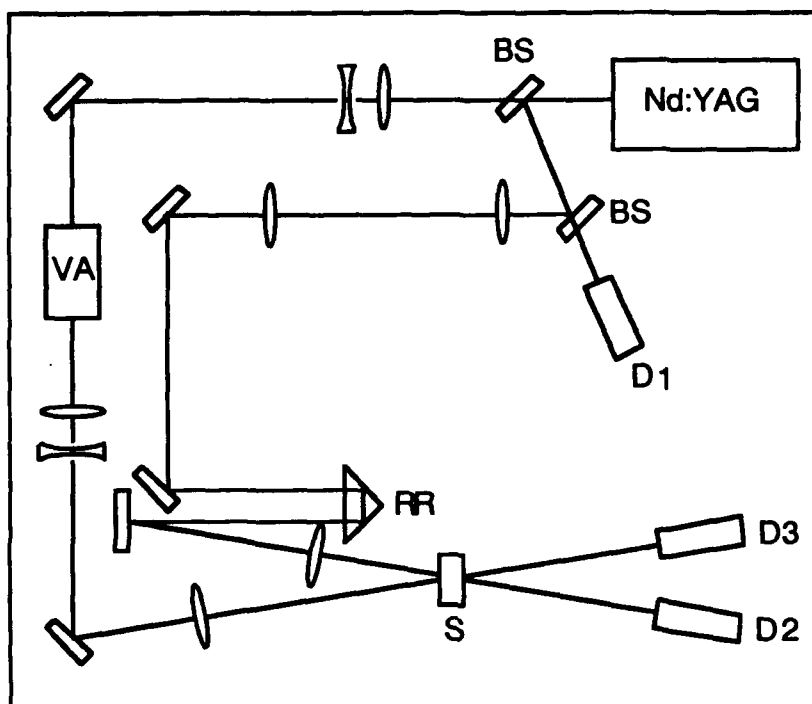


FIGURE 2

Fig.2: Beamline. Pulses of 532 nm light produced by the Nd:YAG laser are split into three beams using uncoated beamsplitters (BS). The sampling detector (D1) measures incident energy on a pulse-by-pulse basis. Detectors D2 and D3 measure probe and pump throughput energies, respectively. A retroreflector (RR) mounted on a translation stage controls probe arrival time. The beam energy is controlled with non-deviating, high power variable attenuators (VA), and the beams are brought to final focus in the sample (S) using a variety of positive and negative focal length UV-grade fused silica lenses and high power mirrors.

At optimum delay a peak gain of 9.0 ± 1.4 is observed, and this delay setting is taken as the arbitrary zero delay point in Figure 3. The gain falls steeply on either side of this point, with a FWHM of about 25 ps. As expected, when the probe arrives much earlier than the pump ($\Delta t < -50$ ps) there is no effective interaction between the two and a gain of zero is observed. An interesting effect was seen at delay settings of $\Delta t > 35$ ps in that the gain is actually less than unity. This is to be

expected if enough free carriers remain in the sample to cause attenuation of the probe.

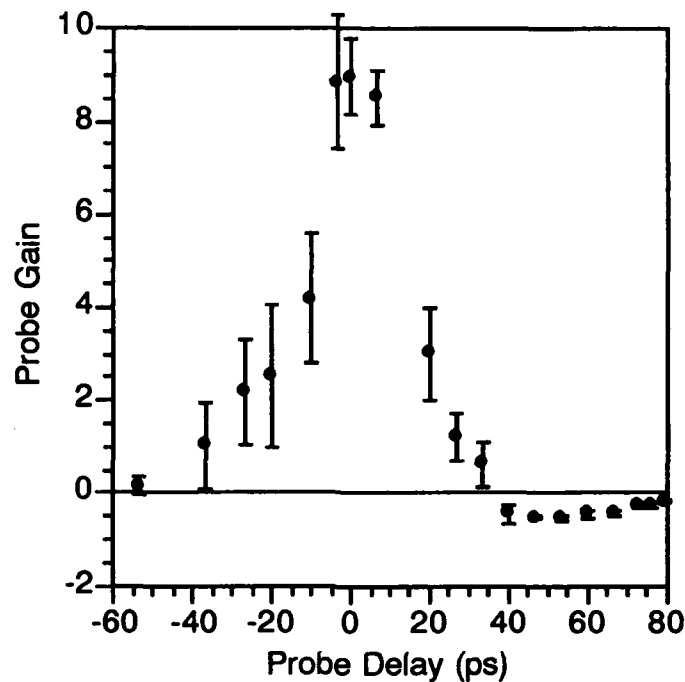


FIGURE 3

Fig. 3: Experimental probe gain as a function of delay between arrival of pump and probe in ZnSe. Positive delay corresponds to probe arriving after pump. The zero is arbitrarily set at peak gain. Peak pump intensity is approximately 3.6 GW/cm².

The code developed was run through a series of probe delay settings in order to obtain a theoretical gain curve. This calculation resulted in a curve with the same basic shape as the experimental data, having a large peak and a region of absorption on one side, as shown in Figure 4. The essential differences between the code plot and experimental data plot are the width of the gain peak (35 ps code versus 25 ps data) and the depth of the absorption valley. The values of β_1 and K_2 that give the best match of the peak gain and maximum absorption to the data are

$$\begin{aligned}\beta_1: & 10 \text{ cm/GW,} \\ K_2: & 3.6 \times 10^{-8} \text{ cm}^3/\text{s.}\end{aligned}$$

A recently published¹⁵ measurement yields a value of $\beta_1 = -8$ cm/GW for ZnSe at 532 nm, comparable in magnitude but differing in sign from our fit value. Substitution of this measured value into our code lowers the peak gain somewhat, but does not otherwise affect the result, illustrating the relative insensitivity of the TET effect in ZnSe to the precise value of β_1 .

Note that the code predicts peak gain will occur when the probe pulse arrives about 10 ps before the pump pulse. With a suitable diagnostic (such as a streak camera) this prediction can be tested.

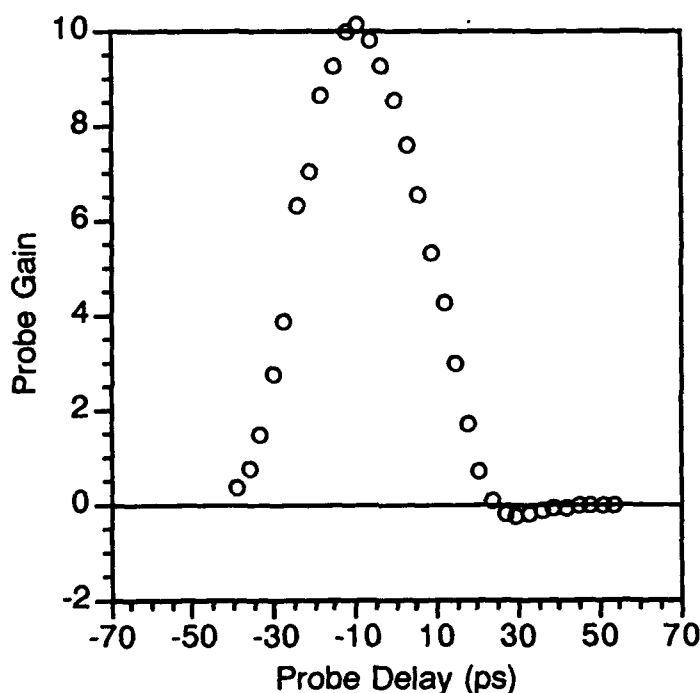


FIGURE 4

Fig. 4: Theoretical probe gain as a function of delay between arrival of pump and probe in ZnSe. Positive delay corresponds to probe arriving after pump. The zero corresponds to true simultaneous arrival of pump and probe. Appropriate experimental parameters were used in performing the calculations.

The experiment is repeated with the probe polarization rotated 90 degrees with respect to the pump. No gain is detected and absorption at the level seen in the

parallel polarization case is found (Figure 5). We believe the primary interaction is simple linear absorption of probe light by pump-produced free carriers. We are testing this model by implementing the code described earlier with appropriate modifications to obtain a second estimate of K_2 for comparison with that reported above. Note the long tail in the recombination, indicative of the slow, linear recombination process neglected in the model.

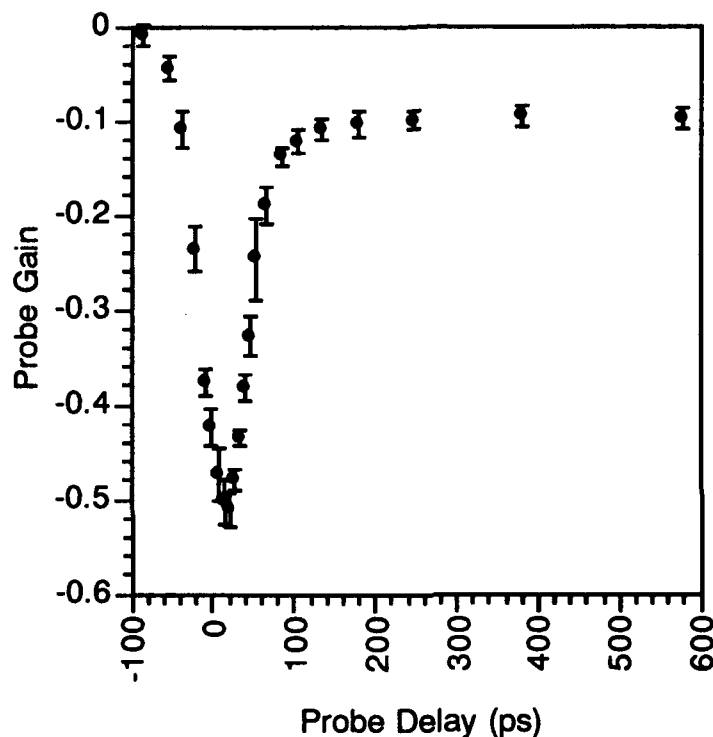


FIGURE 5

Fig. 5: Experimental probe gain as a function of delay between arrival of pump and probe in ZnSe with cross polarized pump and probe. Peak pump intensity is approximately 3.8 GW/cm^2 .

Another study in which ZnS replaces ZnSe is illustrated in Figure 6 below. In this experiment the pump and probe are both vertically polarized and possess the same intensities and energies utilized in the ZnSe study of Figure 3. We observe much lower gain and somewhat greater absorption. The data suggest that loss contributions dominate gain effects in ZnS. We plan to conduct a cross-polarized

pump-probe study to further illuminate the differences between ZnS and ZnSe and to build a numerical model of ZnS.

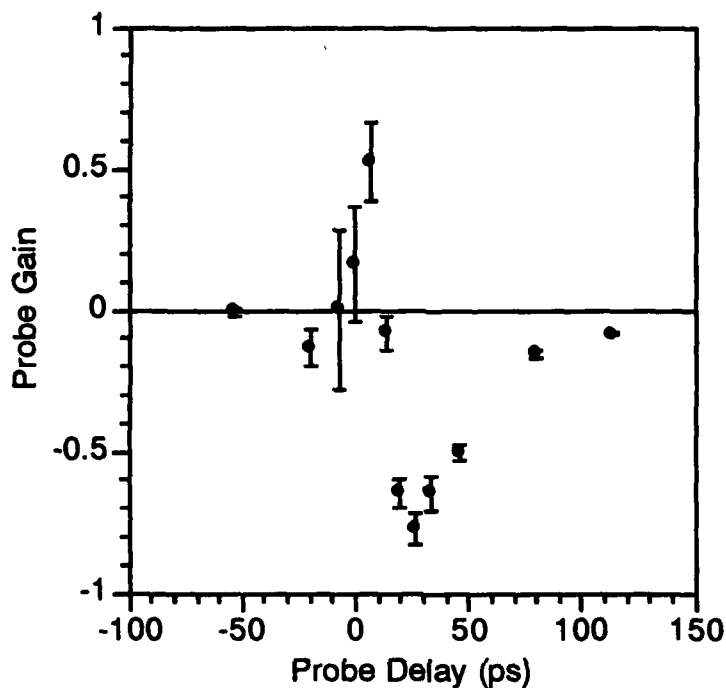


FIGURE 6

Fig. 6: Experimental probe gain as a function of delay between arrival of pump and probe in ZnS. Positive delay corresponds to probe arriving after pump. The zero is at the same zero position set in figure 3. Peak pump intensity is approximately 3.6 GW/cm².

CONCLUSIONS

Under the proper conditions significant energy transfer occurs between pump and probe beams interacting in ZnSe, a wide gap II-VI semiconductor. To a lesser extent energy transfer also occurs in ZnS. Gain in the probe beam varies with pump intensity and relative arrival time between pump and probe. Reasonably good agreement between the code results and the data provide support for modeling the beam coupling interaction with the transient energy transfer theory.

In particular, the values of β_1 and K_2 that gave the best match in peak gain and absorption as a function of probe delay were quite reasonable, the data of reference 15 notwithstanding. However, differences between the data and the theory are significant enough to cause us to re-examine the assumptions and improve the model. In particular, the difference in the width of the gain curves and depth of the absorption tails in Figures 3 and 4 indicate the presence of some systematic error. One possible source is the unverified assumption that the pulses are Gaussian in time. Another is the assumption in the code that the two pulses are fully coherent throughout their lengths. Future work with both code and diagnostics will examine these possibilities. In addition, we plan to extend the numerical modeling to include ZnS and ZnO, and to include diffusion of free carriers. Finally, we hope to expand the code into two dimensions to fully model the beam intensity profile and account for diffraction effects within the crystal.

To date the TET process in wide gap semiconductors has been utilized to study the materials themselves. Applications of the process in a photonic device or process remains to be developed. We believe that interesting applications may be found when the full three dimensional interaction is fully understood. Preliminary studies have shown that the probe is significantly modified spatially as well as in energy in the high gain configuration. In particular we find the amplified probe is deflected somewhat from its original propagation direction, and that its beam quality appears improved. We are presently engaged in conducting these studies.

REFERENCES

1. V. L. Vinetskii, N. V. Kukhtarev, E. N. Sal'kova and L. G. Sukhoverkhova, *Sov. J. Quantum Electron.*, **10**, 684 (1980).
2. V. L. Vinetskii, N. V. Kukhtarev and M. S. Soskin, *Sov. J. Quantum Electron.*, **7**, 230 (1977).
3. G. C. Valley and A. L. Smirl, *IEEE J. Quantum Electron.*, **QE-21**, 488 (1985).
4. See for example J. Feinberg in Optical Phase Conjugation, R. A. Fischer, ed., Academic Press, NY, 1983, p.421.
5. See for example N. W. Ashcroft and N. D. Mermin, Solid State Physics, Holt, Reinhart and Winston, NY, 1976, Ch.1.

6. D. W. Vahey, *J. Appl. Phys.*, **46**, 3610 (1975).
7. Adapted from Ref. 3.
8. Two-Six Inc., Saxonburg, PA 16056.
9. Measured.
10. Related to the nonlinear refractive index n_2 by $\beta_1 = (\omega/c)\gamma$, where $\gamma(\eta^2/W) = (40\pi/cn)n_2(\text{esu})$. Value of n_2 used is taken to be twice that of CdS based on an empirical relationship from N. L. Boling, A. J. Glass and A. Owyong, *IEEE J. Quantum Electron.*, **QE-14**, 601 (1978).
11. E. W. Van Stryland, et. al., *Opt. Engr.*, **24**, 613 (1985).
12. Calculation based on R. K. Jain and M. B. Klein in Optical Phase Conjugation, R. A. Fischer, ed., Academic Press, NY, 1983, p.335.
13. Calculation based on Ref. 6 and mobility and effective mass data from M. A. Omar, Elementary Solid State Physics: Principles and Applications, Addison-Wesley, NY, 1976, Ch.6.
14. Measured; uncertainty +100% -50%.
15. Mansoor Sheik-Bahae, et al., *IEEE JQE* **26**, 760 (1990).

New Optical Switch Stemming from Dipole-dipole Interactions in Dense Media

Michael E. Crenshaw*, Dr., Michael Scalora, Dr., and Charles M. Bowden, Dr.

Weapons Sciences Directorate, AMSMI-RD-WS
Research, Development, and Engineering Center
U. S. Army Missile Command
Redstone Arsenal, AL 35898-5248

I. INTRODUCTION

In dense media, of densities such that there are many atoms (or other effectively two-level systems) within a cubic resonance wavelength, induced near dipole-dipole (NDD) interactions can cause an excitation-dependent renormalization of the resonance frequency.[1-5] Studies of local NDD effects have focused on the steady-state, or stationary, limit in which optical fields vary slowly in an induced-dipole dephasing time or on the time scale of superradiation. In this limit, the field is taken to be constant and it is possible to derive solutions of the equations of motion which show that NDD interactions may cause bistability that is intrinsic to the material and does not require an external feedback mechanism, such as an optical cavity.[1-3] Optical bistability has important applications for optical switching and digital optical memory elements in optical computing and data processing.[6] This phenomenon forms the basis for devices, such as the optical counterpart of the transistor, which constitute the building blocks for all optical light control by light digital optical computing.

In this report, and in a recently published short summary of our work,[7] we describe novel results obtained for the response of a thin film of material, containing or composed of a dense collection of two-level systems (two-level atoms, bound excitons at donor sites, etc.), to resonant excitation by an ultrashort pulse (a coherent optical pulse whose temporal duration is much less than an induced-dipole dephasing time). In such a case, the time development of the applied field is important and the field cannot be taken as constant, as it is in the stationary limit. We find that the induced inversion in the population of the two-level systems that remains in the material after an ultrashort optical pulse has passed has a nearly step-function response to the peak value of the time-varying optical field. The population of the two-level systems, initially in the ground state is always returned to the ground state when the peak field strength is less than the effective dipole moment per unit volume. However, when the peak field strength is in a small region about the effective dipole moment per unit volume, the population is transferred to the excited state and remains in the excited state indefinitely. This result is exceptionally robust with respect to perturbations in parameters, such as variations in pulse shape and width. This nearly step-function response of the system to the field strength constitutes a new and unique optical switch, which, unlike optical bistability,[6] does not have hysteresis and is independent of dissipation from incoherent effects, *i. e.* dephasing and spontaneous relaxation. This can lead to unique, ultrafast, ultra-energy efficient and highly cost effective optical switches and logic devices for application in all optical light control by light optical data processing and optical computing. All-optical digital systems offer the possibility of very high integrity systems, wide bandwidth, high speed, bulk processing, and elimination of optical-to-electronic interfaces.

In the next section, the equations of motion for dense collections of two-level systems, resonantly

excited by ultrashort coherent optical fields, are presented. These equations, derived by Bowden and coworkers,[1, 2, 4] incorporate correction terms into the well-known optical Bloch equations.[8] These correction terms, representing contributions to the field experienced by a dipole from nearby dipoles, often cannot be ignored in condensed matter. These correction terms also appear due to the Lorentz local field correction, but for different reasons and conditions.[3-5] Section 3 presents results of our study of the transient dynamics of dense collections of two-level systems excited by time-dependent coherent optical pulses. No such analysis has previously appeared, except for a short summary of our work.[7] These results, predominately numerical due to the nonlinearity of the equations of motion, indicate a step-function-like response of the system to the peak field strength of an optical pulse. Section 4 discusses possible applications for this response, including optical switching, optical logic, mensuration and material evaluation, and coherent optical pumping. Following the summary, the Appendix presents a previously unreported analytical solution to the equations of motion in a special case.

II. MODEL

The two-level system in quantum optics is a model for two quantum mechanical energy eigenstates connected by a dipole transition. Examples of systems that can be modeled effectively as a two-level system in the optical regime include electronic transitions in atoms, vibrational energy levels in molecules, excitons near the band edge in semiconductors, and laser action.

The optical Bloch equations, which follow from the time-dependent Schrödinger equation for a two-level system, govern the evolution of the system variables in an ensemble of non-interacting two-level systems.[10, 8] Recently, these equations have been extended to include dipole-dipole interactions between two-level systems separated by distances much less than the resonance wavelength.[1-5] For an ensemble of two-level systems interacting, via a dipole transition, with an externally applied, classical, time-varying, coherent field $\frac{1}{2}E(t)[\exp(-i\omega t) + \text{c.c.}]$, the optical Bloch equations, in the rotating-wave approximation and extended to include near dipole-dipole (NDD) interactions in dense media, can be written as[1-5]

$$\frac{\partial u}{\partial t'} = (\Delta - \epsilon w)\tau_p v, \quad (1)$$

$$\frac{\partial v}{\partial t'} = -(\Delta - \epsilon w)\tau_p u + \Omega\tau_p w, \quad (2)$$

$$\frac{\partial w}{\partial t'} = -\Omega\tau_p v, \quad (3)$$

where the system variables u , v , and w are ensemble averages. In writing Eqs. (1)-(3), we have assumed that the pulse is ultrashort, that is, much shorter than the induced-dipole dephasing time T_2 , the population-decay time T_1 , and the superradiation time τ_R , but slowly-varying on the time scale of an optical period. Differentiation is with respect to a dimensionless scaled time $t' = t/\tau_p$, where τ_p is a characteristic time, typically a scaling factor for the temporal width of the pulse. The strength of the NDD interaction is given by $\epsilon = 4\pi\mu^2 N/3\hbar$, where μ is the matrix element of the dipole transition, N is the number of two-level systems per unit volume (number density), and \hbar is Planck's constant divided by 2π . [4] A related quantity, the effective dipole moment per unit volume $\mu_v = 4\pi\mu N/3$ is defined for later use. The instantaneous Rabi frequency $\Omega(t) = \mu E(t)/\hbar$ is the field written in units of frequency. The transition frequency ω_0 is the separation of the energy eigenstates divided by \hbar and $\Delta = \omega - \omega_0$ is the detuning of the field carrier frequency ω from the transition frequency ω_0 . Throughout this paper, the field carrier frequency is assumed to be at exact resonance with the transition frequency ($\omega = \omega_0$), except where explicitly stated otherwise.

The extended Bloch equations, Eqs. (1)–(3), like the ordinary optical Bloch equations, describe the trajectory of a pseudovector, whose components are u , v , and w , on a unit sphere, $u^2 + v^2 + w^2 = 1$. The inversion w represents the difference between the population in the excited state and the population in the ground state. A medium is inverted if there is more population in the excited state than the ground state ($w > 0$). If all of the two-level systems are in the excited state then the medium is in the excited state (fully inverted) which is represented by $u = 0$, $v = 0$, and $w = +1$. Likewise, the medium is in the ground state ($u = 0$, $v = 0$, $w = -1$) if all of the two-level systems are in the ground state. The system variables u and v are proportional to the components of the nonlinear polarization in-phase and in-quadrature with the exciting field. The propagation of a laser field through the medium is described by the extended optical Bloch equations and Maxwell's wave equation, coupled via the nonlinear polarization. While the present work lays the foundation for future studies incorporating propagation effects, we limit the present discussion to the local response by assuming that the film thickness in the propagation direction is much smaller than a resonance wavelength.

III. RESULTS

The inversion is constant, or asymptotically approaches a constant value, after a time-dependent field has passed through the medium ($\Omega(t \rightarrow \infty) = 0$), Eq. (3). In the absence of NDD interactions, an ultrashort resonant optical pulse leaves the medium with an inversion[8]

$$w = -\cos \Theta \quad (4)$$

that depends only on the dimensionless temporal pulse area

$$\Theta = \int_{-\infty}^{\infty} \Omega(t) dt = \frac{\mu}{\hbar} \int_{-\infty}^{\infty} E(t) dt, \quad (5)$$

where we have assumed ground-state initial conditions. Therefore, an ultrashort resonant optical pulse with a pulse area of π (a so-called π -pulse) can be used for coherent optical pumping of a rarefied medium, fully inverting the population. In the process, an amount of energy equal to $n\hbar\omega_0$ is transferred from the field to the medium, where $n = VN$ is the number of two-level systems and V is the irradiated volume. A 2π -pulse returns the population of the two-level systems to the ground state, with no permanent transfer of energy to the medium. A fully inverted medium is often referred to as an amplifying medium because systems in the excited state have stored energy that, under appropriate conditions, can be deposited into an optical field. A medium in the ground state can only extract energy from an optical field and is frequently called an absorbing medium.

To our knowledge, no closed-form solution of the nonlinear differential equations (1)–(3) exists for fields of arbitrary temporal profile in the presence of NDD interactions. Therefore, we numerically solved the extended Bloch equations (1)–(3), using ground-state initial conditions, for the inversion that remains in the dense medium of two-level systems after an ultrashort optical pulse has passed. We find that the asymptotic limit of the inversion (final inversion) has profoundly different characteristics in dense media than the well-known solution, described in the preceding paragraph, in rarefied media.

The final inversion in a dense medium as a function of the peak field strength E_0 (normalized by the effective dipole moment per unit volume μ_v , which is held constant) is shown in Fig. 1 for a Gaussian pulse, $E_0 \exp(-t^2/\tau_p^2)$, in Fig. 2 for a hyperbolic secant pulse, $E_0 \text{sech}(t/\tau_p)$, and in Fig. 3 for a Lorentzian pulse, $E_0/(1 + (t/\tau_p)^2)$. (Note that these curves do not represent a continuous evolution of any system variable, but, rather, each point on the curve corresponds to the outcome of exciting the medium with a pulse that has the peak field strength E_0 .) In these figures, typical of results obtained over a wide range of parameters, there are step-like transitions in the final inversion as a function of

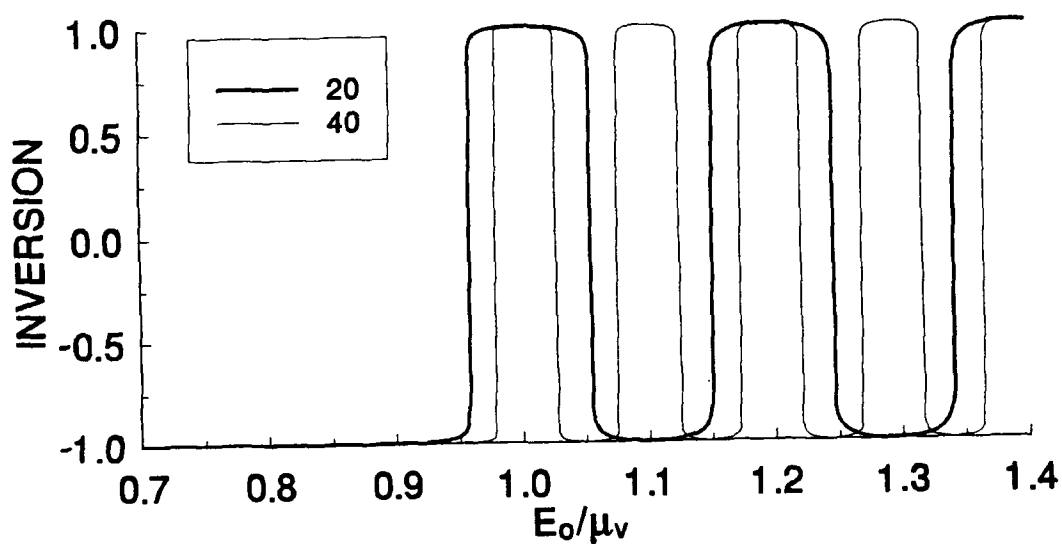


FIG. 1. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort Gaussian pulse for $\epsilon\tau_p = 20$ and for $\epsilon\tau_p = 40$.

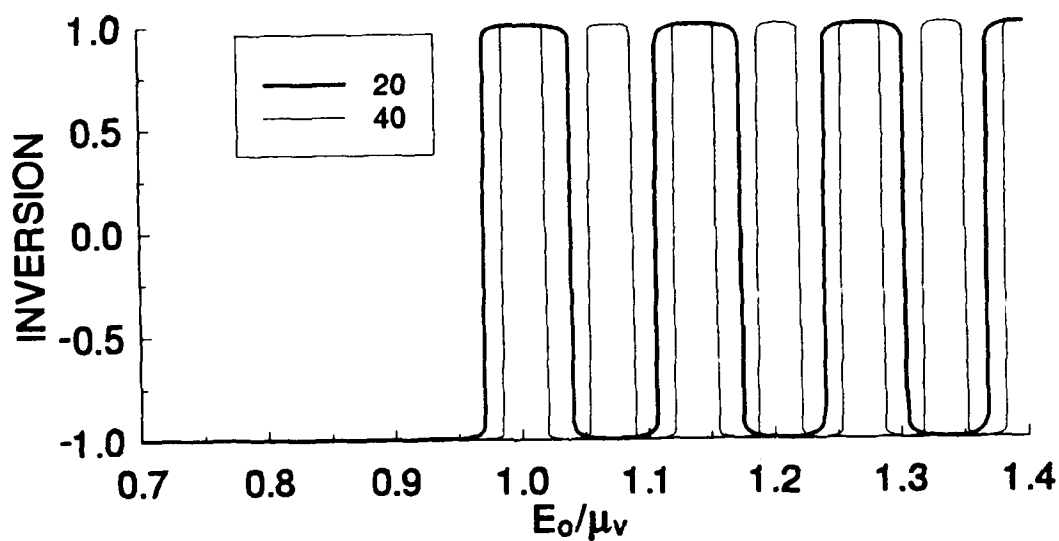


FIG. 2. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort hyperbolic secant pulse for $\epsilon\tau_p = 20$ and for $\epsilon\tau_p = 40$.

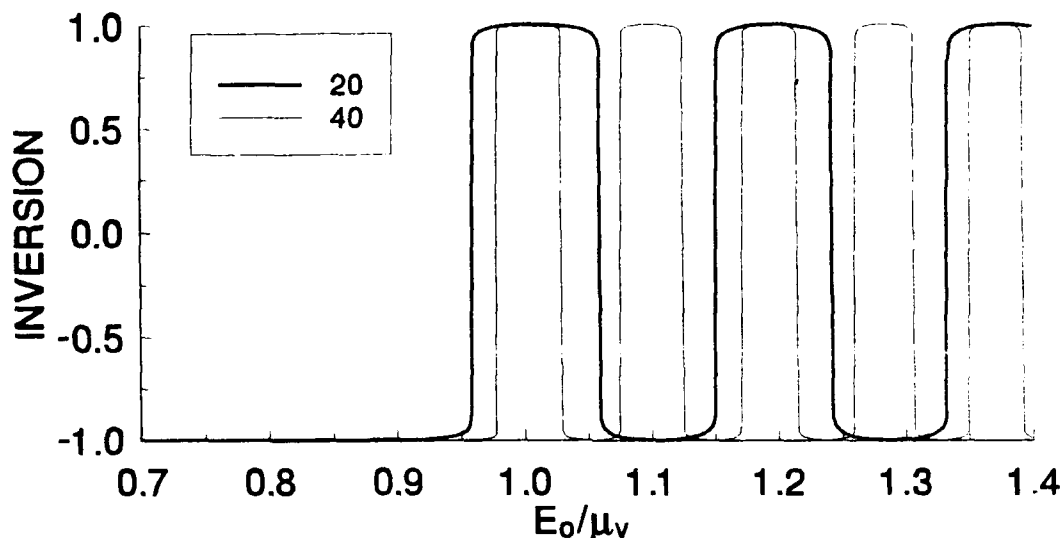


FIG. 3. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort Lorentzian pulse for $\epsilon\tau_p = 20$ and for $\epsilon\tau_p = 40$.

the peak field strength. In each case, the system is left in a fully inverted condition, $w = +1$, when $E_0/\mu_v = 1$. About the point $E_0/\mu_v = 1$ is a region of peak field strength for which the dense medium is left fully, or nearly fully, inverted. Within this region, an amount of energy approximately equal to $n\hbar\omega_0$ is absorbed from the pulse and stored in the medium. The medium is left in an amplifying condition and this response can be considered as analogous to the response of a rarefied medium to the celebrated π -pulse. For peak field strengths on the weak side of this region, the transient field leaves the dense medium in the ground state, in a manner analogous to a 2π -pulse in a rarefied medium. For peak field strengths on the strong side of this region, the final state of the medium alternates between absorbing (ground state) and amplifying (inverted) as a function of peak field strength. Accordingly, we can consider the response of the final inversion to the peak field strength to be approximately a translated square wave with the first half-cycle approximately centered at $E_0 = \mu_v$. Although the pulse area is well-defined for any specific pulse shape and dipole moment, it should be noted that the above results in dense media are independent of the pulse area because the temporal width and shape of the pulse, the number density, and the dipole moment are parameters.

We show in the Appendix that the condition $E_0/\mu_v = (1 + 1/\epsilon^2\tau_p^2)^{1/2}$ defines a value of the peak field strength for which an optical pulse with a hyperbolic secant shape fully inverts the dense medium. Similar results were obtained computationally for Gaussian and Lorentzian shaped pulses with Fig. 4 showing that the value of E_0/μ_v at which the first instance of full inversion occurs can be much greater than unity if $\epsilon\tau_p$ is small, but rapidly converges to unity for $\epsilon\tau_p > 1$. In general, we find that, with increasing $\epsilon\tau_p$, the value of the peak field strength at which the first instance of full inversion occurs approaches μ_v , the inversion region becomes better centered about this point, and the transitions between absorbing and amplifying states become sharper and more periodic. Accordingly, the qualitative results are limited to $\epsilon\tau > 1$.

The above results were presented in the context of simple, symmetric pulse shapes at exact resonance in order to maintain clarity in the exposition. Further numerical tests indicate that the important features of our results are remarkably robust. Figures 5-7, typical of these results, show that the abrupt

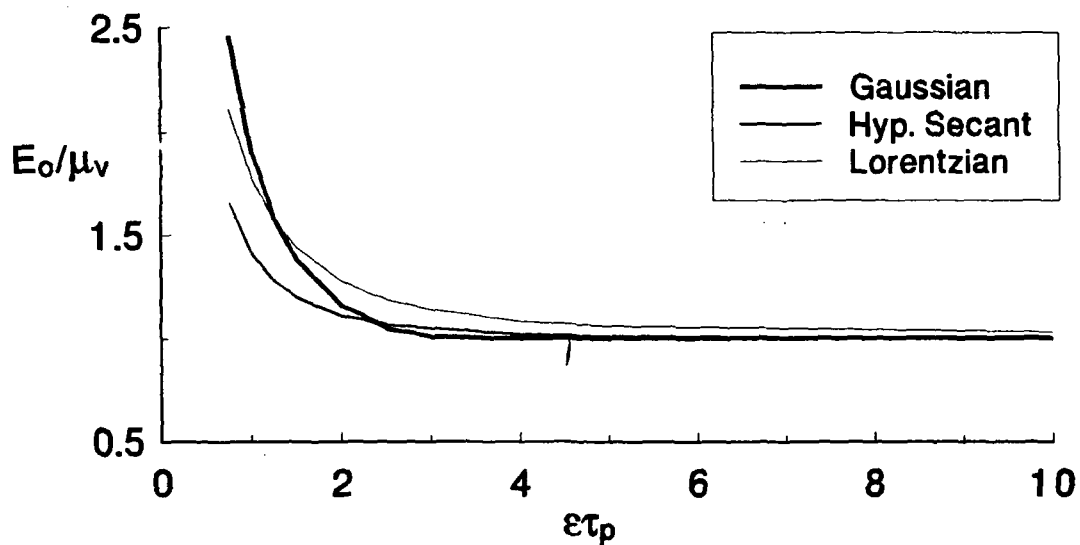


FIG. 4. The value of E_0/μ_v , at the first occurrence of complete inversion, as a function of $\epsilon\tau_p$ for a Gaussian, a hyperbolic secant, and a Lorentzian pulse.

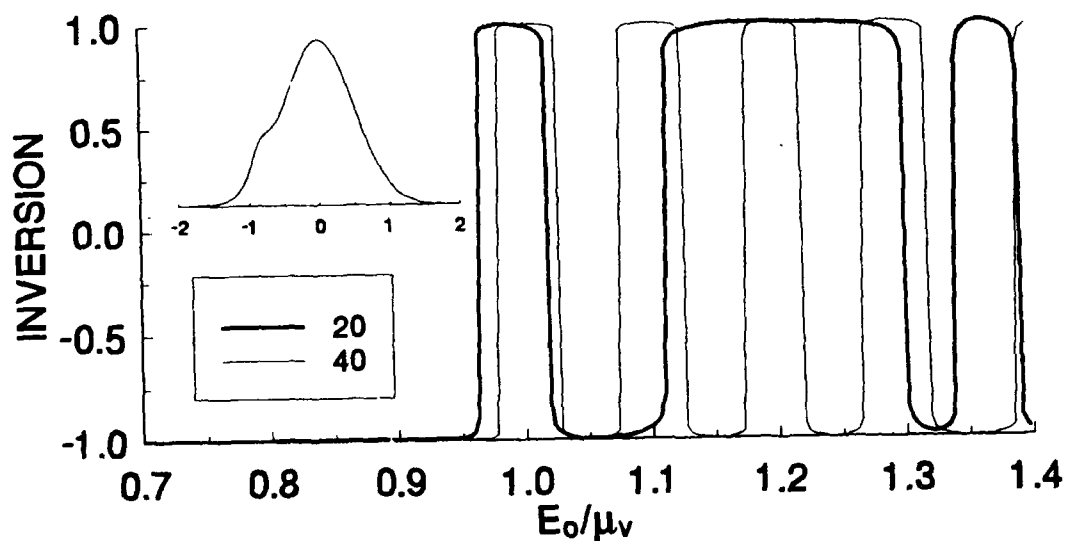


FIG. 5. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort compound Gaussian pulse, $E_0(\exp(-t'^2) + 0.143\exp(-25(t' + 0.875)^2))$ for $\epsilon\tau_p \approx 20$ and $\epsilon\tau_p = 40$. The inset shows the temporal profile of the pulse.

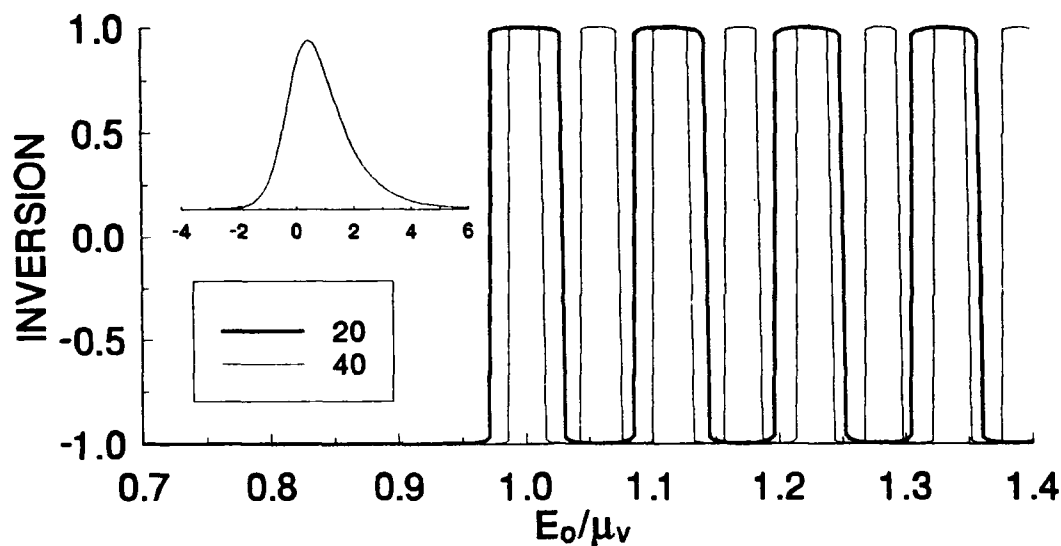


FIG. 6. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort skewed hyperbolic secant pulse, $0.877E_0 \exp(0.5t') \text{sech}(t')$ for $\epsilon\tau_p = 20$ and $\epsilon\tau_p = 40$. The inset shows the temporal profile of the pulse.

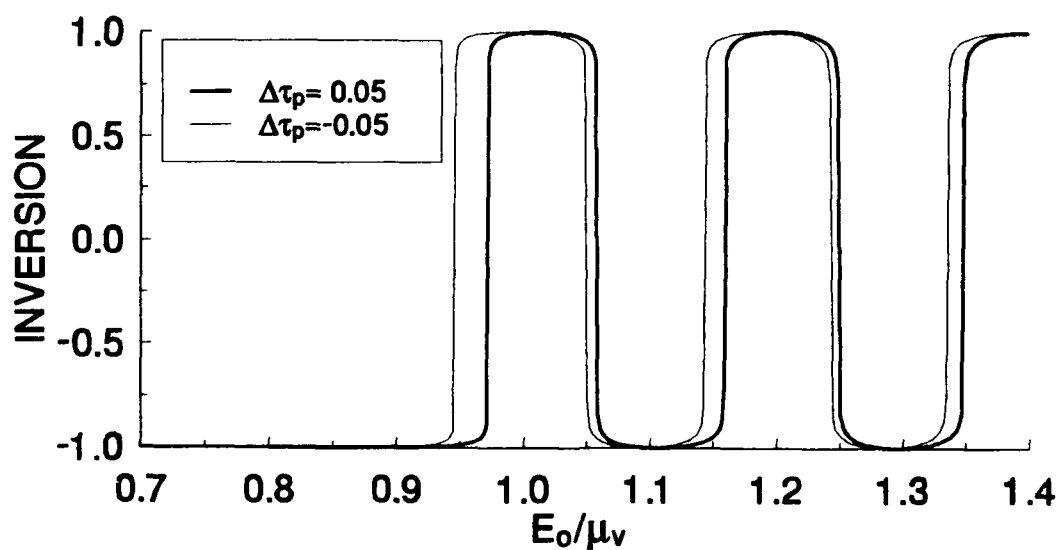


FIG. 7. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort Gaussian pulse ($\epsilon\tau_p = 20$) detuned from exact resonance by $\Delta\tau_p = +0.05$ and $\Delta\tau_p = -0.05$.

transition to a flat inversion region near $E_0 = \mu_v$ persists in the presence of pulse asymmetry (providing the applied field is not highly irregular, such as having multiple local maxima) and moderate detuning from resonance.

As long as the applied field is exactly resonant with the transition frequency, the results obtained for the inversion in a medium that is initially in the excited state $w = +1$ are equivalent to the results obtained for a medium that is initially in the ground state $w = -1$ under change of sign in w . Therefore, a medium in the excited state can be re-inverted by a pulse with $E_0/\mu_v = 1$, returning the medium to the ground state. Detuning the applied field from resonance breaks this inversion symmetry. The final inversion as a function of E_0/μ_v , where the field carrier frequency is detuned to the high frequency side of the transition frequency, is shown in Fig. 8 for both ground state and excited state initial conditions. Figure 9 shows the equivalent results with a negative detuning. Figure 8 shows that, by detuning the field carrier frequency from resonance, it is possible to construct an ultrashort coherent optical pulse such that the medium will be left with an inversion near -1 after the pulse has passed, independent of whether the initial condition is the ground state or the excited state. Likewise, Fig. 9 shows that, with a negative detuning from resonance, it is possible to select a peak field strength such that the inversion is left near the excited state, again independent of whether the initial condition is the ground state or the excited state.

IV. APPLICATIONS

A. OPTICAL SWITCHING AND OPTICAL LOGIC

The step-like transitions in the final inversion as a function of peak field strength suggest that a new kind of optical switch may be devised.[11] The excited state can be interpreted as an "on" state and the ground state can be regarded as the "off" state. The state of the switch can be probed by an ultrashort pulse, with the probe becoming amplified if the switch is "on" or experiencing absorption if the switch is "off".

The basic switching element comprises a thin film of material, typically condensed matter, containing or composed of a dense collection of two-level systems. The population of the two-level systems is assumed to be originally in the ground state due to decay processes, not included in Eqs. (1)-(3), that occur on time scales of the order of the induced-dipole decay time T_2 . The population can be transferred to the excited state by illuminating the thin film with an ultrashort resonant coherent optical field with a peak field strength approximately equal to the effective dipole moment per unit volume, $E_0 = \mu_v$. In the process, an amount of energy equal to $n\hbar\omega_0$ is absorbed from the field and stored in the medium. A second resonant pulse, again with a peak field strength approximately equal to the effective dipole moment per unit volume, is then used to probe the medium. If the energy of the probe, upon exiting the medium, is greater than the energy input by $n\hbar\omega_0$ then the medium was in the excited state. However, if the energy is reduced by $n\hbar\omega_0$ then the medium was in the ground state. Note that, for exactly resonant pulses, the process of probing the state inverts, or re-inverts, the medium. The basic operation of this switch can be considered to be a digital optical flip-flop. The switch may also be used as a pulse height window discriminator, or comparator, in which pulses with peak field strengths between a lower threshold and an upper threshold change the state of the switch.[12]

There is a clear application of the optical switch just described to a two-input AND gate with the inputs being identical pulses with peak field strengths of $\mu_v/2$. Assume that the medium is initially in the ground state. The medium remains in the ground state if neither of the pulse inputs pass through the medium. If either of the pulse inputs pass through the medium, then the medium is partially excited, but returns to the ground state. Only if both pulse inputs are present, simultaneously, will the optical field have the necessary field strength, $E_0 = \mu_v$, to transport the population of the two-level systems to the excited state. Similarly, both inputs must be present in order to re-invert a medium

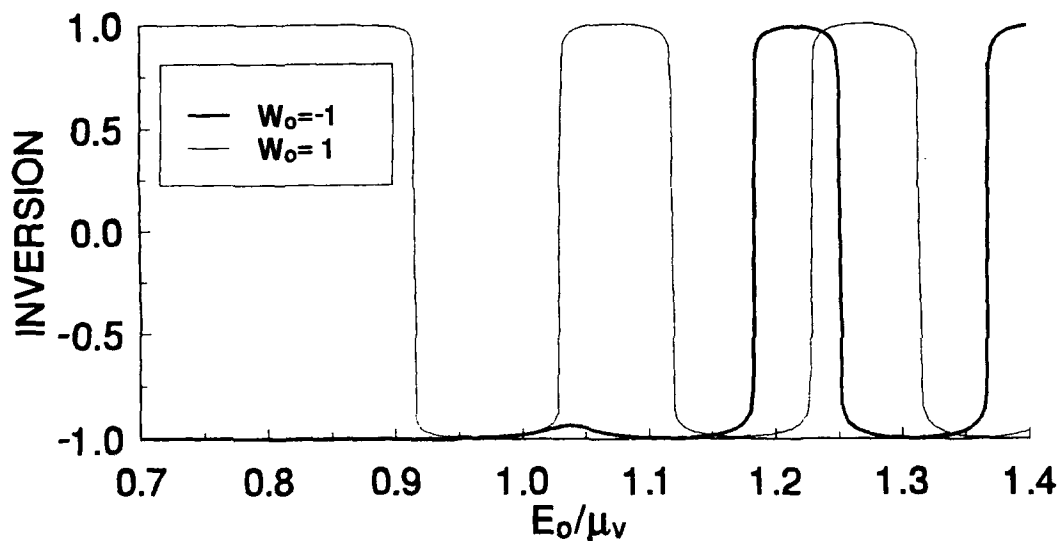


FIG. 8. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort Gaussian pulse ($\epsilon\tau_p = 20$) detuned from exact resonance by $\Delta\tau_p = +0.2$ for ground state and excited state initial conditions.

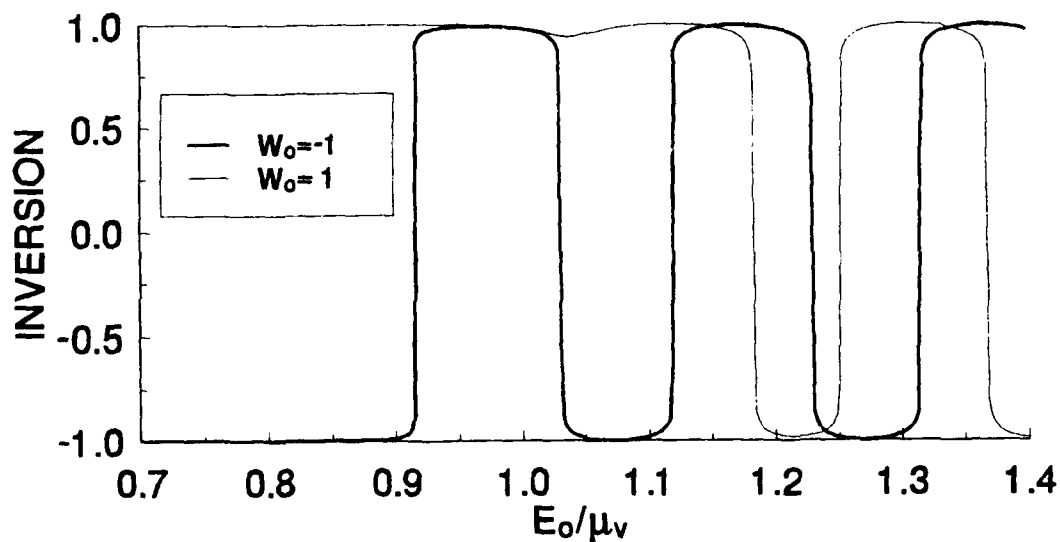


FIG. 9. The final inversion, as a function of the normalized peak field strength E_0/μ_v , due to an ultrashort Gaussian pulse ($\epsilon\tau_p = 20$) detuned from exact resonance by $\Delta\tau_p = -0.2$ for ground state and excited state initial conditions.

that is initially in the excited state. The input/output table for the two-input AND gate is shown in Table 1 for ground state initial conditions and in Table 2 for excited state initial conditions. A NOT operation, inverting or re-inverting the medium, is performed by an optical pulse with a peak field strength $E_0 = \mu_v$. Therefore, the two-input AND operation can be inverted by preceding or following the inputs with an optical pulse such that $E_0 = \mu_v$, creating a two-input NAND gate.

Additional logic operations may be performed using a reference-plus-inputs scheme. Let δ denote the width, in terms of the variation of peak field strength, of the full inversion region about μ_v and let the inputs be pulses, each with a peak field strength of $\delta/2$. A NOR operation can be performed by combining the pulse inputs with a reference pulse that has a peak field strength of $\mu_v + \delta/4$. It is assumed that the reference pulse and the pulse inputs have the same temporal development. If neither of the inputs is present, then the reference pulse inverts, or re-inverts, the medium. If either or both of the inputs are incident on the medium simultaneously with the reference pulse then the peak field strength of the combined field is outside the full inversion region about $E_0 = \mu_v$. In this case, the medium is not inverted or re-inverted, but is returned to its initial state. The input/output table for the NOR gate is shown in Table 3 for ground state initial conditions and in Table 4 for excited state initial conditions. Similarly, a NAND operation is performed if the reference pulse has a peak field strength of $\mu_v - \delta/4$, as shown in the input/output tables, Tables 5 and 6. It is possible to extend this reference-plus-inputs scheme to OR and AND operations, using reference pulses with peak field strengths of $\mu_v - 3\delta/4$ and $\mu_v - 5\delta/4$, respectively. However, these latter two operations may not be necessary because preceding or following the NAND or NOR logic operations with the reference pulse inverts the logic operation.

It is generally necessary to be able to assign a specific state to a switch, as well as to be able to change the state. For instance, we may want to load a memory register with a specific bit pattern, or to clear (reset) a register, without knowing the previous contents. We demonstrated in Fig. 8 that, if the carrier frequency is detuned to the high frequency side of the transition frequency, a pulse with a peak field strength just below μ_v can be used to place the medium in the ground state, regardless of whether the initial state is the ground state or the excited state, thereby resetting the switch. Likewise, a pulse can be used to set the switch, placing the medium in the excited state, regardless of whether the initial state is the ground state or the excited state, by detuning to the low frequency side of the transition frequency, as was demonstrated in Fig. 9. While it is not necessary to know the previous state of the switch in order to put it in the desired state, this information can be deduced (query and reset) by determining whether the pulse has gained energy, lost energy, or has no change in energy.

The energy required to excite the medium is clearly $n\hbar\omega_0$. However, the total energy of the switching pulse will generally be greater. The kind of switching times and energies required can be calculated from the condition $\Omega_0 = \epsilon$. Restrictions on the parameters that must be satisfied are $\epsilon \ll \omega_0$ and $\tau_p \ll T_1$ and T_2 . Further, our results indicate that it is reasonable to require $\epsilon\tau_p > 1$. Assuming a pulse width of one picosecond, a dipole moment of $10ea_0$, and an active device area of $1 \mu\text{m}^2$, we determine the total pulse energy to be of the order of a femtojoule at optical frequencies. It is significant that the energy transfer between the medium and the fields is coherent. This allows very high repetition rates without heat damage to the material (heating can be a significant problem in absorptive optical bistability, limiting bit rates).

B. MENSURATION AND MATERIAL EVALUATION

The step-like transitions between the ground state and the excited state about the point $E_0/\mu_v = \Omega_0/\epsilon = 1$ constitute a unique and unambiguous signature of NDD interaction. Materials in which these effects might be observed include oxygen ions in $\text{KCl}:\text{O}_2^-$ and bound I_2 excitons at donor sites in CdS single crystals.[13, 14] Of special significance is the fact that these step-like transitions can be used to develop sensitive techniques for measuring the strength of the NDD coupling parameter ϵ . One can obtain the value of ϵ by simply increasing the peak field strength of an ultrashort pulse from shot-to-

Table 1: AND GATE ($w_0 = -1$)

Input A	Input B	A + B	w
0	0	0	-1
$\mu_V/2$	0	$\mu_V/2$	-1
0	$\mu_V/2$	$\mu_V/2$	-1
$\mu_V/2$	$\mu_V/2$	μ_V	1

Table 2: AND GATE ($w_0 = +1$)

Input A	Input B	A + B	w
0	0	0	1
$\mu_V/2$	0	$\mu_V/2$	1
0	$\mu_V/2$	$\mu_V/2$	1
$\mu_V/2$	$\mu_V/2$	μ_V	-1

Table 3: NOR GATE ($w_0 = -1$)

Ref.	A	B	A+B+R	w
$\mu_V+\delta/4$	0	0	$\mu_V+\delta/4$	1
$\mu_V+\delta/4$	$\delta/2$	0	$\mu_V+3\delta/4$	-1
$\mu_V+\delta/4$	0	$\delta/2$	$\mu_V+3\delta/4$	-1
$\mu_V+\delta/4$	$\delta/2$	$\delta/2$	$\mu_V+5\delta/4$	-1

Table 4: NOR GATE ($w_0 = +1$)

Ref.	A	B	A+B+R	w
$\mu_V+\delta/4$	0	0	$\mu_V+\delta/4$	-1
$\mu_V+\delta/4$	$\delta/2$	0	$\mu_V+3\delta/4$	1
$\mu_V+\delta/4$	0	$\delta/2$	$\mu_V+3\delta/4$	1
$\mu_V+\delta/4$	$\delta/2$	$\delta/2$	$\mu_V+5\delta/4$	1

Table 5: NAND GATE ($w_0 = -1$)

Ref.	A	B	A+B+R	w
$\mu_V-\delta/4$	0	0	$\mu_V-\delta/4$	1
$\mu_V-\delta/4$	$\delta/2$	0	$\mu_V+\delta/4$	1
$\mu_V-\delta/4$	0	$\delta/2$	$\mu_V+\delta/4$	1
$\mu_V-\delta/4$	$\delta/2$	$\delta/2$	$\mu_V+3\delta/4$	-1

Table 6: NAND GATE ($w_0 = +1$)

Ref.	A	B	A+B+R	w
$\mu_V-\delta/4$	0	0	$\mu_V-\delta/4$	-1
$\mu_V-\delta/4$	$\delta/2$	0	$\mu_V+\delta/4$	-1
$\mu_V-\delta/4$	0	$\delta/2$	$\mu_V+\delta/4$	-1
$\mu_V-\delta/4$	$\delta/2$	$\delta/2$	$\mu_V+3\delta/4$	1

shot (*e. g.*, by removing neutral density filters) and locating the points where the sharp transitions in the state of the medium occur by the change in the transmitted energy. The midpoint between the peak field strengths at these transitions is the effective dipole moment per unit volume μ_v . Likewise, the midpoint between the peak Rabi frequencies $\Omega_0 = \mu E_0/\hbar$ at these transitions is the NDD parameter ϵ . The precision of the measurement depends on τ_p , because this parameter determines the width of the full inversion region. Measurements of the frequency renormalization in dense media have been performed using the reflectivity of an interface between a sapphire window and a dense potassium vapor.[9] A portion of the observed frequency shift was attributed to collisions.

C. OPTICAL PUMPING

Coherent optical pumping refers to coherent excitation of a medium of two-level systems which transfers population from the ground state to the excited state. The well-known π -pulse method of coherent optical pumping is not valid in dense media. However, excitation with an ultrashort pulse that has a peak field strength within the full inversion region about the effective dipole moment per unit volume constitutes a method for coherent optical pumping of dense media.[15] While the width of the full inversion region depends on parameters such as pulse width, pulse shape, dipole moment, and density, this method of optical inversion is less sensitive to perturbations in the parameters than the π -pulse method where the inversion falls off as the cosine of the pulse area. Typical perturbations that might occur are pulses with a spatial profile or which experience mild absorption, or the pulse width, pulse shape, dipole moment, or density may not be known to a high degree of accuracy or may be difficult to control. Applications for coherent optical pumping in dense media include laser gain media, superfluorescent lasing, and coherent spectroscopy. The method may also be used to re-invert the media, transferring population from the excited state to the ground state.

V. SUMMARY

We have identified a unique and unambiguous signature of NDD interaction in dense collections of two-level atoms: An ultrashort pulse leaves the population of a dense collection of two-level systems in the ground state or the excited state with a step-function-like dependence on the peak field strength. This result can be used to measure the NDD coupling parameter and for a coherent optical pumping mechanism if propagation effects and medium relaxation times can be neglected, *i. e.* for thin films and short optical pulses. A fast, energy-efficient optical switching mechanism has been proposed. This optical switch can form the basis for optical logic devices for application in all-optical-light-control-by-light optical data processing and computing.

ACKNOWLEDGMENTS

Michael Scalora is a National Research Council Research Associate.

APPENDIX

TIME DEVELOPMENT OF THE INVERSION

Some interesting results for dense collections of two-level systems, excited by ultrashort pulses, can be obtained by studying the time development, as well as the asymptotic limit, of the inversion. The time development of the inversion, with E_0 as a parameter, is shown in Fig. A1 for an applied field, expressed in units of frequency, given by

$$\Omega(t) = \Omega_0 \text{sech } t / \tau_p = \frac{\mu E_0}{\hbar} \text{sech } t', \quad (\text{A1})$$

where the applied field is exactly resonant with the transition frequency. The pulse area, Eq. 5, for this field is $\Omega_0 \tau_p \pi$.

The curve in Fig. A1 that represents transfer of population from the ground state to the excited state suggests that the extended Bloch equations (1)-(3) may have a solution in the form

$$u = a \text{sech } t' \quad (\text{A2})$$

$$v = b \text{sech } t' \quad (\text{A3})$$

$$w = c \tanh t'. \quad (\text{A4})$$

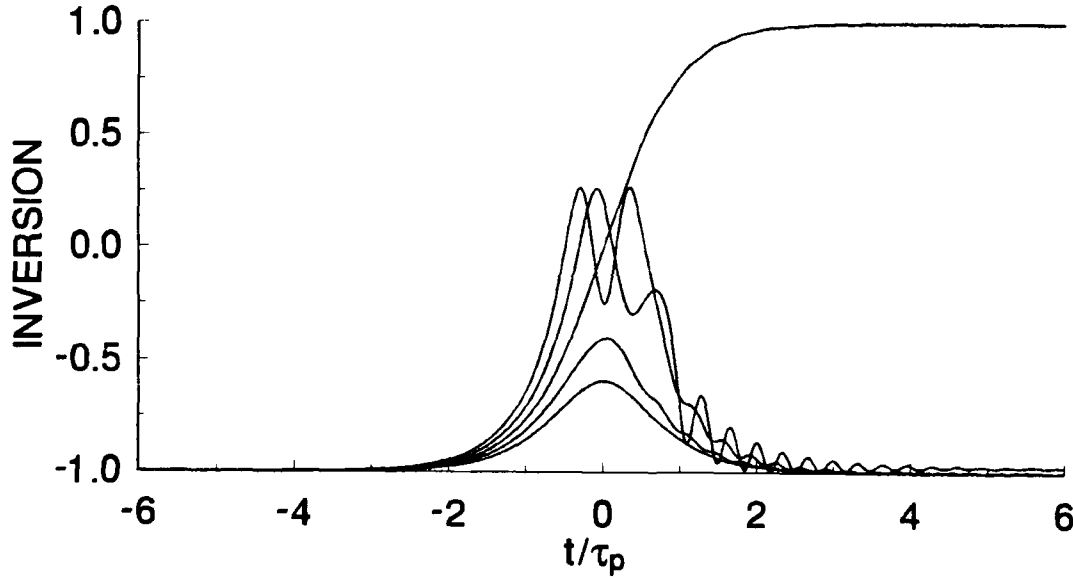


FIG. A1. The time development of the inversion, due to an ultrashort hyperbolic secant pulse, for E_0 equal to $0.8\mu_V$, $0.9\mu_V$, $1.0\mu_V$, $1.1\mu_V$, and $1.2\mu_V$.

Substituting the *ansatz* (A1)-(A4) into the extended Bloch equations (1)-(3) (with $\Delta = 0$) yields one set of equations for the derivatives of the system variables, while taking the derivatives of Eqs. (A2)-(A4) with respect to t' yields another. Equating these two sets of equations, we obtain

$$a = -\frac{\Omega_0^2 \tau_p^2 - 1}{\epsilon \Omega_0 \tau_p^2}, \quad b = \pm \sqrt{\frac{\Omega_0^2 \tau_p^2 - 1}{\epsilon^2 \Omega_0^2 \tau_p^4}}, \quad c = \pm \sqrt{\frac{\Omega_0^2}{\epsilon^2} - \frac{1}{\epsilon^2 \tau_p^2}}.$$

If the medium was in the ground state ($u_0 = 0$, $v_0 = 0$, $w_0 = -1$) (or the excited state ($u_0 = 0$, $v_0 = 0$, $w_0 = +1$)) at some time in the indefinite past ($t \rightarrow -\infty$), then Eqs. (A2)-(A4) are a solution to the extended optical Bloch equations (1)-(3) for the field (A1) if, and only if,

$$\frac{\Omega_0}{\epsilon} = \frac{E_0}{\mu_v} = \sqrt{1 + \frac{1}{\epsilon^2 \tau_p^2}} \quad (\text{A5})$$

at exact resonance. Consistency with the initial conditions clearly requires $\Omega_0 = \epsilon$ ($E_0 = \mu_v$) in the limit $\epsilon^2 \tau_p^2 \rightarrow \infty$. Further, near-equality of E_0 and μ_v holds for $\epsilon \tau_p > 1$, in accordance with our numerical results. In the opposite limit of small $\epsilon \tau_p$, the medium is inverted by a hyperbolic secant pulse when Ω_0 is approximately $1/\tau_p$. In this limit, the pulse area $\Omega_0 \tau_p \pi$ approaches π , *natura non facit saltum*.

The extended Bloch equations (1)–(3) can be written in the form

$$\ddot{w} - \frac{\dot{\Omega}}{\Omega} \dot{w} + \Omega^2 \tau_p^2 w + \Omega \epsilon \tau_p^2 w \left(1 - w^2 - \frac{\dot{w}^2}{\Omega^2 \tau_p^2} \right)^{1/2} = 0 \quad (\text{A6})$$

using the identity $u^2 + v^2 + w^2 = 1$. Equation (A6) contains dynamical frequency and damping terms, leading to the damped oscillations seen in Fig. A1, as well as dynamical anharmonicities.

- [1] Y. Ben-Aryeh, C. M. Bowden, and J. C. Englund, *Phys. Rev. A* **34**, 3917 (1986).
- [2] Y. Ben-Aryeh and C. M. Bowden, *Opt. Commun.* **59**, 224 (1986).
- [3] M. G. Benedict, V. A. Malyshev, E. D. Trifonov, and A. I. Zaitsev, *Phys. Rev. A* **43**, 3845 (1991).
- [4] C. M. Bowden, *Near Dipole-Dipole Interaction Effects in Quantum and Nonlinear Optics*, (International Conference on Quantum Optics, Hyderabad, 1991, unpublished).
- [5] F. Friedberg, S. R. Hartmann, and J. T. Manasseh, *Phys. Rev. A* **40**, 2446 (1989).
- [6] H. M. Gibbs, *Optical Bistability: Controlling Light with Light*, (Academic Press, New York, 1985).
- [7] M. E. Crenshaw, M. Scalora, and C. M. Bowden, *Phys. Rev. Lett.* **68**, 911 (1992).
- [8] L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms*, (Wiley, New York, 1975).
- [9] J. J. Maki, M. S. Malcuit, J. E. Sipe, and R. W. Boyd, *Phys. Rev. Lett.* **67**, 972 (1991).
- [10] R. P. Feynman, F. L. Vernon, and R. W. Hellwarth, *J. Appl. Phys.* **28**, 49 (1957).
- [11] M. E. Crenshaw, M. Scalora, and C. M. Bowden, *Active Element for Optical Signal Processing and Optical Computing*, MICOM Invention Disclosure, AMPC4760 (1992).
- [12] H. V. Malmstadt, C. G. Enke, and S. R. Crouch, *Electronics and Instrumentation for Scientists* (Benjamin/Cummings, Menlo Park, 1981), Chap. 4.
- [13] M. Dagenais and W. F. Sharfin, *Appl. Phys. Lett.* **45**, 210 (1984).
- [14] M. S. Malcuit, J. J. Maki, D. J. Simkin, and R. W. Boyd, *Phys. Rev. Lett.* **59**, 1189 (1987).
- [15] M. E. Crenshaw, C. M. Bowden, and M. Scalora, *Method for Optically Inverting a Dense Collection of Two-level Systems*, MICOM Invention Disclosure, AMPC4761 (1992).

DASCHKA, RUZGA, PETREANU, KUMAR

Ceramic Matrix Fiber Reinforced Composites

Paul Daschka, Mr., Richard Ruzga, Mr., John Petreanu, Mr.
and *Ashok Kumar, Dr.

U. S. Army Construction Engineering Research Laboratory
P.O. Box 9005
Champaign, Illinois 61826-9005

INTRODUCTION

Polymeric and metallic coatings have typically been used to protect steel from corrosion. There are some problems associated with both of these types of coatings. Polymeric organic coatings such as vinyls have low resistance to erosion corrosion. Environmental Protection Agency (EPA) regulations require that volatile organic compounds (VOC) be reduced. In addition, red lead oxide was used in the past as a primer for polymeric coatings and is now considered a hazardous waste. Removal and disposal of lead based paint (LBP) requires special equipment and procedures which can increase the cost of surface clean-up and preparation by a factor of 2 to 3 over non-lead based paint.

Metallic coatings such as zinc, aluminum, and stainless steel can be applied by fusion processes (welding) or by thermal spray processes. Production rates of these processes are low, polymeric sealers must be added if the thermal spray process is used, and stainless steel or brass coatings may cause galvanic corrosion problems.

Many of these problems may be solved or avoided through the use of ceramic coatings such as enamels. These coatings have no VOC emissions and are environmentally safe. However, ceramic compounds are known to be brittle and prone to chipping. Surface preparation for ceramic coatings can include thermal treatments for the in-situ dissolution of corrosion products. This may also be applicable to lead compounds which can potentially reduce the cost of LBP abatement.

The objective of this study was to develop and evaluate

ceramic matrix, fiber reinforced composites for application in repair of corrosion damage to construction steel or other metallic components used by the U.S. Army.

Several methods are available for repair and rehabilitation of painted steel structures, depending on operating conditions and environment. The most commonly used non-fusion (non welding) coatings are epoxies reinforced with ceramic or metallic particles. Although these materials have been successfully used in specific situations, the predictability of their performance is low. The bond strength between the epoxy and metallic substrate depends on the surface condition and, with time, the repair compounds may fail due to mechanical fatigue of the bond interface due to corrosion, cavitation, or erosion. The use of reinforced epoxy materials can help mitigate galvanic corrosion that results from using a metal repair material dissimilar to the original component composition. Porcelain enamel coatings may be reliable and effective repair materials for damage in metallic parts based on their chemical bonding, resistance to cavitation and erosion damages, and ability to be used in portable thermal spray coatings units (Reference 1 to 3).

The enamel composition developed and investigated in this work was designed for the repair of cavitation and erosion damage without the need to remove existing surface oxidation. Calcium silicate fibers were added to the composition to increase the fracture toughness and increase the overall damage resistance. The cavitation and erosion resistance of the enamel was compared to a reinforced epoxy material (Reference 4 and 5). In addition the glassy compositions have potential for application in in-situ vitrification of lead based paint on structures.

CALCIUM SILICATE FIBERS

Although silicon carbide fibers have excellent strength and toughening properties, their oxidative instabilities at high temperatures could prove unsuitable for glass or enamel composites. The fibers may reduce the glass or enamel. Calcium silicate fibers may be used in these situations. Although the strength of the calcium silicate fiber (approximately 1.2MPa) is much less than SiC fiber, it does have quite a few advantages. The fiber has a high melting point (1540 degree C), high corrosion resistance, and because of its inertness at elevated temperatures, can be used in an oxygen environment without harming the matrix. Furthermore, because these fibers

occur naturally, they are relatively cheap.

Wollastonite is the naturally occurring calcium metasilicate, CaSiO_3 . It is the only commercially available pure white mineral that is wholly acicular. Table 1 contains physical properties of wollastonite fibers. Typical aspect ratios range from 3:1 to 20:1. The average fiber diameter is 3.5 microns and has a coefficient of thermal expansion of $6.5 \times 10^{-6}/^\circ\text{C}$. Although there has been limited research on calcium silicate used as a reinforcing fiber, wollastonite has been used traditionally in the ceramic, paint, and paper industries. Pure grades of wollastonite are used in minor quantities in rubber, plastics, and other manufactured items.

DESCRIPTION OF TESTING MATERIALS

The substrate material used in this work was A36 steel. This steel is often specified for use in construction applications. The linear thermal expansion coefficient of A36 steel is $6.5 \times 10^{-6}/^\circ\text{C}$. The composition ranges and limits (weight percent) are as follows:

Carbon	< 0.29
Manganese	0.80 - 1.20
Copper	> 0.20

The epoxy control material used was a commercially available reinforced epoxy material, Ceramic S-Metal, produced by Belzona Inc. This epoxy material has been widely used in industry as a sacrificial coating for the repair of hydraulic turbines and mechanical equipment. The epoxy is made by mixing a base and a solidifier in a 3:1 ratio. A thin coat was applied to the A36 steel substrate and allowed to cure for 24 hours.

Wollastonite (calcium metasilicate) fibers were used to reinforce the enamel coating. This fiber has a high melting point, 1540°C , good corrosion resistance, low cost, and resistance to being reduced at high temperatures. In addition, the small fiber size and diameter (3.5μ) makes it suitable for spray applications.

The physical properties are illustrated in Table 1, while the enamel composition used for this study is illustrated in Table 2. It was derived by adjusting several batch compositions to get a composition that exhibited a strong chemical bond with the A36 steel, did not fail during thermal cycling, and was compatible with the calcium silicate fibers. The calculated

DASCHKA, RUZGA, PETREANU, KUMAR

thermal expansion value of this composition is $2.5 \times 10^{-5}/^{\circ}\text{C}$.

Table 1.

Physical Properties of Wollastonite Fibers

Property	Value
Appearance	Brilliant White
Shape	Acicular
Molecular Weight (g/mol)	116
Specific Gravity (g/cm ³)	2.90
pH (aqueous solution)	9.9
Water Solubility (gm/100cc)	0.0095
Density (lbs/solid gallons)	24.2
Mohs Hardness	4.5
Coefficient of Expansion (mm/mm°C)	6.5×10^{-6}
Melting Point (°C)	1540
Aspect Ratio (typical)	20:1
Surface Area (BET, m ² /gm)	0.88

Table 2.

Glass Enamel Composition

(Network Forming)	Silicon Oxide	=	44.68 g
(Network Forming)	Boron Oxide	=	18.46 g
(Network Modifying)	Sodium Oxide	=	10.49 g
(Network Modifying)	Potassium Oxide	=	4.40 g
(Network Modifying)	Cobalt Oxide	=	2.21 g
(Network Modifying)	Calcium Fluoride	=	11.42 g
(Intermediate Oxide)	Aluminum Oxide	=	4.77 g
(Intermediate Oxide)	Manganese Oxide	=	3.57 g
			<hr/> 100.00 g

ENAMELING PROCESS

Standard enameling procedures were used in this work for preparing samples for testing. This procedure consists of three major steps: enamel/fiber preparation, substrate preparation, and enamel application to the substrate in a furnace.

ENAMEL/FIBER PREPARATION

The enamel batch was prepared from the amounts of raw materials contained in Table 2. After the batch is made from the raw materials, the enamel must be fritted. Fritting glass and enamels involves melting the raw materials to form a homogeneous molten batch, and then quenching the molten material in water until it returns to room temperature. This process produces shattered fragments with the desired composition. Fritting is important to assure quality control by making the molten batch homogeneous and to evolve gases and impurities which might effect the coating quality. If a glass or enamel was melted directly from a raw batch, the uneven mixing would produce a slightly different composition each time. Large shrinkages are also typical for glasses which were not fritted.

The fritting operation used in this work was accomplished by using a conventional microwave oven. Microwave heating is a relatively new technique for processing ceramics. In this work, iron titanate was used as the heating source as follows:

- 1) A large crucible was half filled with the enamel batch, slightly compacted.
- 2) Iron titanate was sprinkled on a heating brick in the microwave oven and the crucible was placed on top of the material with a hollow insulating cylinder around it.
- 3) A thin piece of insulation was placed on top of the crucible, leaving a small opening to allow for escape of gases evolved during the melting.
- 4) The assembly was placed on top of insulating board to prevent the glass oven tray from breaking due to excessive heating.
- 5) The oven door was closed and the heating initiated. The heating was interrupted every 10 minutes to add more of the enamel batch and to stir the molten material.

DASCHKA, RUZGA, PETREANU, KUMAR

6) The crucible was removed from the microwave oven and the molten glass was poured into a water bath. This procedure was done quickly so the enamel would not solidify in the crucible.

7) The microwave oven was allowed to cool before another batch was melted.

The fritted enamel was then ground into a powder to allow for intimate mixing with the fibers. The frit was placed in a closed ball mill cylinder which was 2/3 filled with distilled water and contained 2.5 cm diameter alumina balls. The container was then milled for 12 hours and the frit was dried and separated through a 270 mesh screen.

For the fiber reinforced enamel, the fibers can be added to the uniform frit to make the enamel-fiber matrix. The mixing procedure was as follows:

- 1) The fibers were mixed with the enamel powder.
- 2) A 25% mixture of deionized water by volume was added.
- 3) The pH was raised to 10 to allow for deflocculation.
- 4) The mixture was stirred for 15 minutes.
- 5) The pH was lowered to the original value.
- 6) The mixture was stirred with heat until the water evaporated.

Deflocculation is necessary since the glass tends to agglomerate and does not allow the even mixing of the fibers. Care must be taken when choosing a buffering solution since it may affect the characteristics of the glass.

ENAMEL APPLICATION

The enamel powder was deposited on the cleaned A36 steel samples by the use of a spray gun. A slurry was made by adding the powder to an appropriate amount, 1 gram of powder to 1 ml alcohol, of isopropyl alcohol. A small amount of distilled water was added to increase the sticking of the enamel to the substrate. The slurry was then applied in two passes and allowed to dry. Two coats were applied to the samples, the first was the enamel without fibers, to increase adhesion. Fibers at the interface may impede the bonding between the enamel and the steel. The second coat, applied after the first coat had slowly cooled from heat treatment, contained fibers in 5, 7, and 10 weight percents. Samples were prepared without fiber additions, and these too were composed to two layers to

DASCHKA, RUZGA, PETREANU, KUMAR

achieve similar coating thicknesses as the fiber containing coatings.

The samples were heat treated at 800°C for 15 minutes. For the samples used for laboratory evaluation, the substrates were slow cooled inside the furnace. Gradual cooling helps to relieve stress gradients due to uneven cooling at the surface and enamel bulk. A few samples were removed quickly from the furnace for rapid cooling. A rapid cooling schedule simulates field applications, where furnace cooling is not possible.

Flame spray methods were investigated for field applications for damage repair. The system used was a TeroDyn System 3000, manufactured by Eutectic Corp. It employs an oxy-acetylene flame and includes a forced gas powder feeder to prevent enamel agglomeration. A laboratory test investigating the deposition of the fiberless and with 7% fiber reinforced enamel was conducted and the processing parameters were optimized.

EROSION TESTING

The laboratory erosion test employed has been described by Weber et. al. (Reference 4). The test apparatus consisted of a 20 liter porcelain coated steel container with internal baffles to minimize vortex generation with the slurry and an electric motor driving the stainless steel specimen holder. The specimen holder resembles a four bladed paddle with four specimens mounted 45 degrees to the vertical. The motor was a high torque 220V, 3 phase system with an electronic controller and magnetic digital tachometer to control the speed. The entire system was fixed to maintain unit stability.

The A36 steel 1.25 cm thick plate stock was cut into 3 X 5 cm specimens. A run consisted of 4 samples being tested simultaneously. One sample was the CERHAB without fiber, the second was CERHAB with either 5, 7, or 10% fiber addition, the third was the reinforced epoxy material, and the fourth was an un-coated A36 steel sample used as a control. The specimens were weighed to the nearest 0.1 milligram, mounted into the specimen holder and were rotated at 575 revolutions per minute (RPM) by the motor in the sand/water slurry for a total of 72 hours. The slurry consisted of 1.2 liters of dry sand (250 μ particle size) mixed with 5.8 liters of water. This yielded a 17 volume percent solution. This slurry is two to three times more concentrated than that found in the Ohio River. The apparatus was stopped at hours 1, 5, 10, 24, 48, and 72 and the

specimens were cleaned, dried at 105°C for 1 hour, and weighed. In this way a history of mass loss versus time can be obtained for each of the specimens as in the cavitation test. By comparing these, the relative erosion resistance of each of the materials can be determined.

CAVITATION TESTING

The vibratory apparatus used in this investigation was a Ultratip Labsonic System manufactured by Lab-line. This apparatus produces axial oscillations of the test specimen partially immersed in distilled water. The vibrations are produced by a piezoelectric transducer that is driven by an electronic oscillator and power amplifier (Reference 5). A sound box provided hearing protection during operation.

The specimen was immersed into a 250 ml beaker containing the water. The amplitude and frequency of vibration of the test specimen was controlled and monitored in order to maintain optimal cavitation conditions. The test specimen was cleaned and weighed accurately before each test, and again during periodic interruptions of the test in order to obtain a cumulative mass loss versus time history.

SCANNING ELECTRON MICROSCOPY (SEM)

Both the enamel/steel and epoxy/steel interfaces were examined using scanning electron microscopy (SEM) to qualitatively compare the bonding of the enamel and epoxy. A high resolution ISI DS-130 SEM unit was used for this investigation. A JEOL JSM U-3 SEM unit was used to observe the enamel and epoxy surfaces. An accelerating voltage of 15 keV was used. The interface specimens were prepared using a Buehler Isomet low speed saw with a wavering blade. The samples were mounted using a mixture of Duco cement and conducting graphite. To prevent charging, the specimens were gold plated using a sputter coater operating at 50mA for 30 seconds. The samples were polished with 180, 240, 320, 400, and 600 grit sandpaper, followed by 5.0, 1.0, and 0.3 micron polishing paste.

The enamel coating prepared using the traditional furnace approach proved to be satisfactory. When applied to small steel surfaces, the fiberless enamel did not fail under either cooling schedule. The un-reinforced enamel did fail on a large steel surface (5" X 12") exposed to rapid cooling. This was caused by cohesive failure within the enamel. The enamel did not fail when 7% by weight calcium silicate fibers were added. The

calcium silicate fibers appear to have reinforced the enamel coating to the extent that it did not fail during thermal cycling. Although care was taken during design to keep the coefficient of thermal expansion as low as possible, the value was calculated to be $2.5 \times 10^{-5}/^{\circ}\text{C}$. This value is higher than the $6.5 \times 10^{-6}/^{\circ}\text{C}$ value for A36 steel. The expansions seemed to be matched with respect to the lack of failure encountered.

SEM micrographs were taken of the interface between the coatings and metal substrate to compare qualitatively the bonding of the enamel and the epoxy. The epoxy/metal interface is shown in Figure 1 (Top). The metal is on the left and the reinforced epoxy material is on the right. It is evident that the interface has large gaps running along the two surfaces. It was believed that the main problem with the epoxy was its weak bonding to the metal. However, after close inspection, this may not be entirely true. This large gap is caused by the metal cohesively cracking near the interface. Since an exothermic reaction is involved in making the epoxy, this large gap may be the result of thermal expansion mismatches. The only apparent bonding results from mechanical interlocking along the interface.

In contrast, the enamel/metal interface, Figure 1 (Bottom), reveals intimate contact between the two surfaces. The enamel appears to be wetting the metal completely. In addition, some fragments can be seen in the enamel. These are thought to be oxide products from the metal or other surface defects. The defects are safely pulled away from the interface and bond strength is preserved.

Initial fabrication using Nicalon β -SiC fibers failed because of reactions at elevated temperatures reduced the enamel causing it to fail. The addition of calcium silicate fibers appears to have helped prevent cohesive failure during thermal cycling and increased the erosion resistance of the enamel. It appears that the calcium silicate fibers were not taken into solution, and that crack deflection toughening is the likely mechanism for the improved toughness from thermal shock.

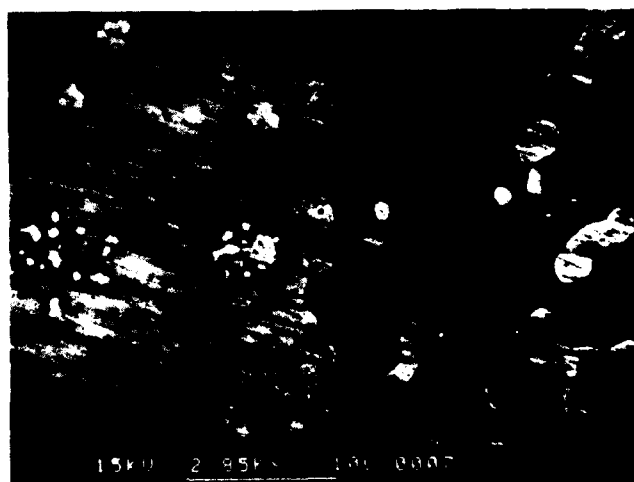
RESULTS OF EROSION AND CAVITATION TESTING

Seven erosion tests were performed, with the samples as stated previously. Two tests for each fiber reinforced sample were performed to correlate the data. A great deal of heat is generated by the unit after 72 hours of operation. The slurry temperature was raised to approximately 70°C by the conclusion



Figure 1. (Top) x450
SEM Micrograph of the Epoxy/Metal Interface.
Mechanical interlocking is the only evident bonding feature.

Figure 1. (Bottom) x2850
SEM Micrographs of the Enamel/Metal Interface.
Enamel-metal interface demonstrating intimate
contact along surfaces



of the test. This is a much higher temperature than which hydraulic equipment is usually subjected, and probably increased the erosion rate of all of the samples. The erosion pattern on the samples were not uniform, the leading edge of each sample was the most eroded. Also the outside corners exhibited more degradation, possibly from the increased velocity of the outside edges.

Three cavitation tests were performed for each coating condition. There were three samples tested in all of the tests: an epoxy coating, a reinforced enamel (5, 7, or 10% fiber addition) coating, and an un-reinforced enamel coating. Time intervals between weighings were 2 minutes during the early portion of the test and five minutes until the conclusion. The results of cavitation tests are given in cumulative weight loss versus time curves. Particular attention was paid to the slopes of the curves. Coatings with less slope had slower cavitation rates and thereby provided better cavitation resistance.

It is interesting that as the weight percent of the fiber reinforcement increases, the resistance to cavitation damage slightly decreases. It appears that there is an optimum (7 percent) fiber weight percent to give maximum performance, as is true with many fiber reinforced coating systems.

Figure 2 (Top), shows the results of the cavitation test performed using the 7% fiber reinforced enamel. The figure clearly indicates that the enamel is much more resistant to cavitation than the reinforced epoxy. This suggests that any added fracture toughness that the fibers give to the enamel has little effect on its resistance to cavitation because the fiberless enamel performs as well as or better than the fiber reinforced enamel.

Figure 2 (Bottom), shows the results of the erosion tests performed using the 7% fiber reinforced enamel. It is clearly seen in both tests, with and without fibers, that the enamels exhibit no noticeable advantage over the reinforced epoxy. However, the 7% fiber enamel performs similarly to the reinforced epoxy while the fiberless enamel does not. The improvement in the erosion and corrosion resistance over the bare A36 steel is attributed to the fibers which are present in the fiber reinforced CERHAB as well as in the reinforced epoxy. Both the reinforced epoxy and the enamels are much more erosion resistant than the un-coated A36 sample.

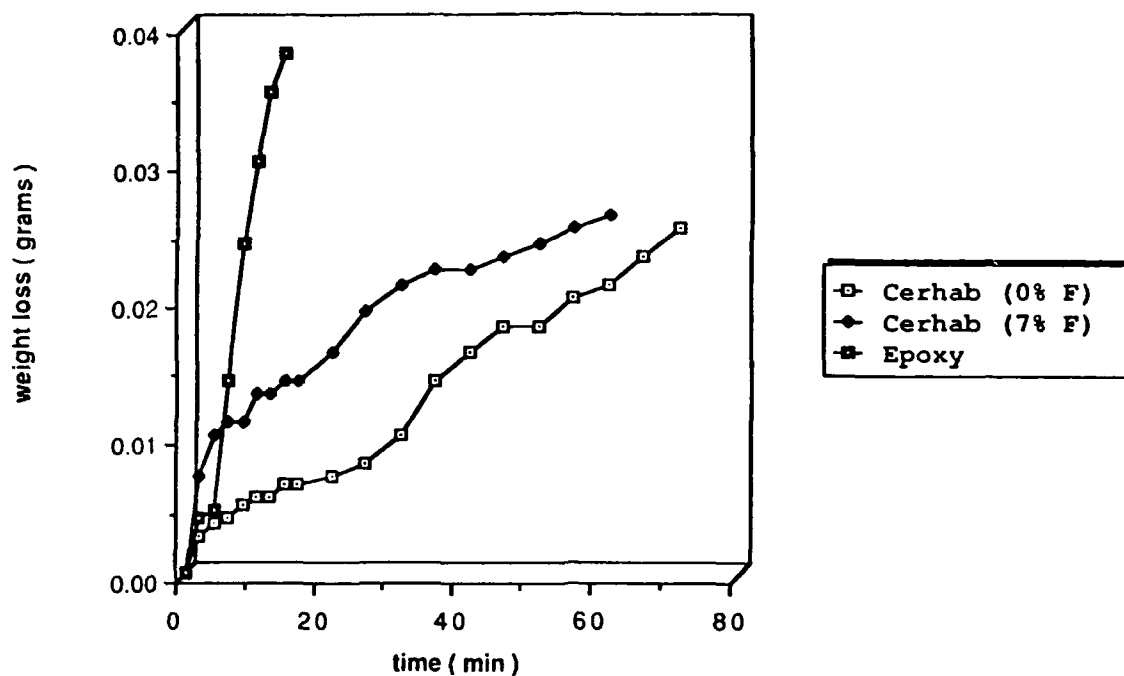
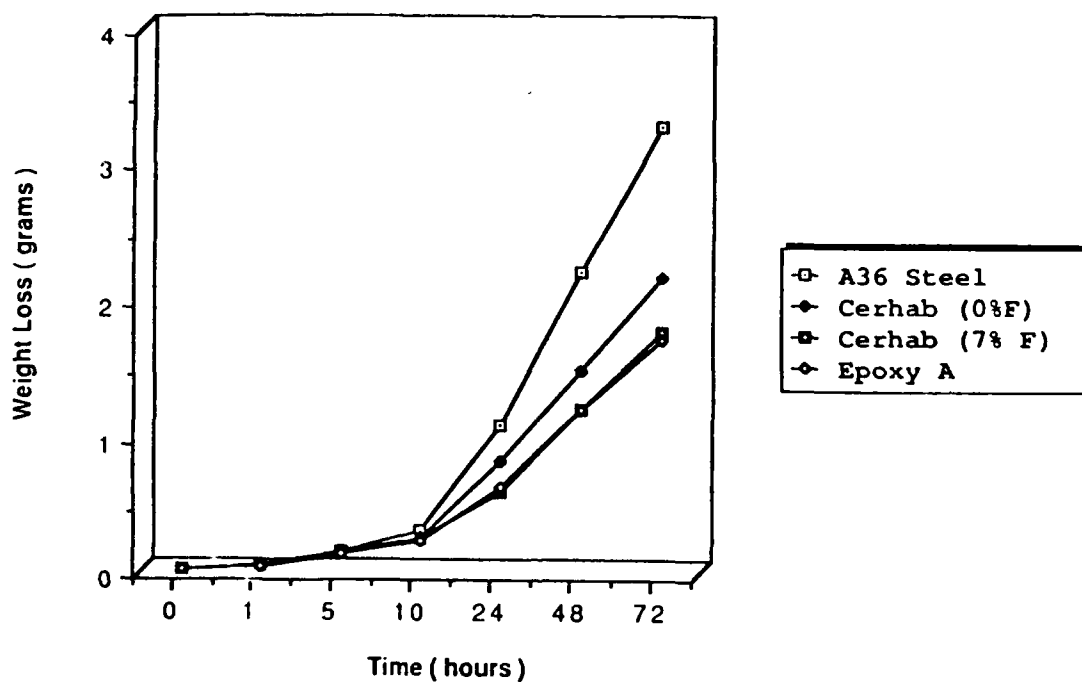


Figure 2. Cavitation test results (Top) and erosion test results (Bottom).



INVESTIGATION OF APPLICATION PROCEDURES

Two samples were sprayed to demonstrate the feasibility of the TeroDyn 3000 flame spray process to deposit fiber reinforced enamel coatings on A36 steel substrates (15 cm X 10 cm X 0.3 cm) in a shop environment. The first test consisted of spraying the fiberless enamel powder with a forced powder delivery system. This was used because the small particle size and shape of the enamel powder can plug the regular feed system. The coating was extremely prone to cracking, therefore after the coating process was complete, a fusing step was added. The coating was fused using a oxy-acetylene torch. The second test consisted of spraying a 7% fiber reinforced enamel on the A36 steel substrate. This test was performed by the Eutectic Corporation and met with mixed success. As is the case with most glass materials, internal stresses developed on cooling from the liquid state. These stresses when superimposed over the stresses produced by the differences in thermal expansion differences caused fracture and debonding when the coating was applied to large surfaces. The coatings were successful when applied to small substrates.

Annealing cycles may be necessary to relieve the stresses in the coating during the cooling stage. The CERHAB coating process would not be applicable in field situations where the substrates are large and require annealing. The annealing process can be accomplished by strap-on heaters. The annealing procedure used in the shop test to produce the CERHAB coatings involved maintaining the coating at 1000°F for two hours, and then cooling at a rate of 122°F/hr until room temperature was reached. It is common in welding repairs to preheat or post heat the welded area, therefore similar equipment could be employed when using ceramic coatings.

A36 platestock, 1.25 cm thick, was cut into 6x6 cm specimens. The squares were painted with lead oxide paint and allowed to cure for up to three weeks. Based on the lead alkali ternary system various compositions were developed to determine the maximum lead oxide solubility. The processing scheme followed was that of the CERHAB composition excluding the fiber reinforcement. The TeroDyn 3000 flame spray system was used to apply the compound to the specimens. Micro X-Ray Spectrometry (XRS) techniques were used to determine the amount of lead oxide taken into solution. Compositional as well as processing parameters are being optimized for maximum lead oxide removal.

The CERHAB compound, though developed for erosion/corrosion repair, appears to have even more potential for use as a coating and abatement technique for lead-based paint on steel structures such as bridges. The glassy composition of CERHAB can dissolve the lead oxide present on the steel structure along with iron oxide in a process called in-situ vitrification. Thermal stress developed in the glass caused by rapid cooling of the glassy coating will help cause cracking followed by chipping and removal of the coating. Alternately, the fiber-reinforced glassy coating can be annealed by heating and cooling at a slow rate, resulting in a fiber-strengthened durable coating. Lead oxide has been used as a primer on steel structures and is considered a hazardous material. The currently used method of removal by sand blasting requires the removed material be contained and that workers be protected from exposure to lead. In addition, the remaining sand mixed with lead oxide is considered a hazardous waste and must be treated before disposal. This can increase the cost of removal of the lead-based paint from steel oxides by a factor of 2 to 3. The silica glassy composition has the potential for use as a lead-based paint abatement and disposal technology.

CONCLUSIONS

- 1) The formulated glass enamel CERHAB provides a strong gap-free chemical bond with the metallic substrate and takes into solution metallic oxides, safely pulling them away from the interface.
- 2) The glass enamel CERHAB shows an increased resistance to cavitation when compared to the epoxy material. This may result from its strong bond to the steel substrate.
- 3) The reinforcement of the enamel CERHAB provided by the calcium silicate fibers reduces the catastrophic fracture upon cooling, and increases the erosion resistance when optimized at 7 percent by weight. It has little effect on the cavitation resistance of the enamel. The major toughening mechanism appears to be crack deflection. Fiber pullout was not observed.
- 4) The fiber reinforced CERHAB can be successfully flame sprayed on small steel substrates. There has been difficulty in flame spraying this compound on large surfaces. Annealing by electrical heaters or heat lamps will be needed on larger component parts.
- 5) The CERHAB enamel displays strong chemical bonding,

DASCHKA, RUZGA, PETREANU, KUMAR

good resistance to cavitation, and the ability to be applied with a flame spray unit (followed by annealing). The silica glassy composition shows potential for application and use as a repair material on metallic components and for lead based paint abatement.

REFERENCES

1. Ashok Kumar, John Bonini, and John Bukowski "Ceramic Fiber Reinforced Glasses for Corrosion Damage Repair" Paper No. 84, Corrosion 89. National Association of Corrosion Engineers, (April 1989).
2. Ashok Kumar, John Bonini, and Paul Daschka "CERHAB for Corrosion Damage Repair" Paper No 479 Corrosion 90. National Association of Corrosion Engineers Las Vegas (April 90).
3. P. Daschka, J. Bonini, R. Ruzga and A. Kumar "Using Ceramic Fiber-Reinforced Glasses for Corrosion, Cavitation and Erosion Damage Repair of Metal Components" USACERL Technical Report M-92, (1992).
4. R. Weber, M. Rigsbee, M. Brandt, T. Race, I. Calace and V. Hock "An evaluation of Metallic Coatings and Test Procedures for Erosion/Corrosion conditions at Lock and Dam Structures" in E. L. Houch. ed. Thermal Spray: Advances in Coatings Technology: Proceedings of the Thermal Spray Conference (ASM International, 1988).
5. Standard Method of Vibratory Cavitation Erosion Test ASTM G32 (American Society for Testing and Materials, (1985).

DE CAMP

MAXIMIZE Automatic Target Recognition Algorithm Development Environment

Brian C. DeCamp
CECOM Night Vision & Electro-Optics Directorate
Ft. Belvoir, VA 22060-5677

I. The Need for Automatic Target Recognition (ATR)

Automation of weapon systems will play a major role in the army of the future. The performance of early-generation automated equipment in the Desert Storm conflict demonstrated the importance of this technology. The development of automatic target recognizers capable of locating and identifying threat vehicles and weapon systems on the battlefield is an area of automation receiving particular attention. Weapon systems with ATR capability will cue the soldier to the presence of possible threats. This technology will help the Army improve lethality of Fire-and-Forget weapons, reduce operator workload in manned weapon systems and improve battlefield management techniques for Identification-Friend-or-Foe. As weapon systems become more complex a greater load is placed on the operator of the system. Aiding or automating the process of finding, identifying and tracking threats in the battlefield reduces the stress placed on the operator therefore improving survivability.

II. The Need for Automatic Target Recognition Evaluation Tools

Nearly all ATR algorithms developed so far are made up of a few common components: interest point location (finding target-like objects in the image), segmentation (finding silhouettes of the target-like objects), feature extraction (measurement of geometric properties of the objects), and recognition (comparison of measured features with a library of features for known targets). Many statistical algorithms have all four of these steps in the decision chain. Other types of algorithms exist which combine two or more algorithm components into a single component. For instance, some algorithms use geometric and infrared signature models of targets to make "Model-Based Algorithms" which combine the steps of features extraction and recognition into a single component.

Since it is not likely that the best algorithm suite will come from any one source, it is necessary to connect algorithm components from different sources and evaluate the performance of the new suites of algorithms to discover better recipes of computer operations for better

DE CAMP

ATR performance. All of these factors multiply the complexity of the process of determining the best algorithm for a mission requirement.

In the past, standard techniques for evaluating these components have not existed and algorithm developers have had to rely on loosely controlled experiments on small data sets and intuition for selecting the best components for their algorithms. This is further complicated by the fact that each of the components may have several custom thresholds or other types of parameters which change the characteristics of the algorithm. It was difficult for algorithm designers to empirically evaluate the effects of these parameters without efficient tools.

While it is important to conduct experiments in a reasonable amount of time, the most difficult problem facing the ATR developer is developing, interfacing, and evaluating the various algorithms being researched. To effectively perform this task one needs a Computer-Aided Software Engineering (CASE) tool which will allow the user to integrate different algorithm components from different sources. This includes integrating algorithms which were developed by competing organizations.

The CASE tool must also be available in the public-domain so that it may be freely distributed to algorithm developers for use during their development cycle. It must allow both symbolic information and other ancillary data, such as the name and coordinates of a target, to be attached to the image. There must be a mechanism in such a tool that allows algorithms to take in custom run-time variables (i.e. thresholds). Since a typical evaluation of an algorithm requires processing tens of thousands of images for statistically significant results, the tool must be able to efficiently coordinate processing in the distributed computer environments of the 1990's.

The concept of reconfigurable image operations is not new to CASE technology. In fact at least three major software products provide this capability. Unfortunately, none of these packages satisfy the requirements of being available in the public-domain, allowing ancillary data to be included in the output of an algorithm component and efficiently distributing processing to a network of computers.

III. MAXIMIZE

The CECOM Night Vision & Electro-Optics Directorate has developed the MAXIMIZE software facility to aid in the development and evaluation of ATR algorithms. This facility has allowed the Army to combine algorithms from different sources and improve the state-of-the-art of the ATR algorithm development process.

One of the most important goals of MAXIMIZE was to provide a mechanism for reconnecting ATR algorithm components by setting standards for the input and output of each of the different components. For instance, every interest point location algorithm makes some estimate of the cartesian coordinates of the center of the object in question. If every programmer stores this information in a different way, the interest point location algorithm cannot

communicate with any other algorithm which needs this information.

Algorithm components written for MAXIMIZE use an object-oriented dialect of the "C" programming language. MAXIMIZE specifies how to create and access "C"-objects which contain the types of information typically found in ATR algorithms, such as the cartesian coordinates of an interest point. These objects may then be passed from one algorithm component to another serving as a communication protocol for standard types of information. The flexibility of object-oriented design also allows application-specific information to be added to these objects as algorithms evolve.

MAXIMIZE consists of an intuitive Graphic User Interface (GUI) which allows the user to select ATR operations from various menus and specify data path connections between these modules as shown in Figure 1. The ATR created can be executed from the GUI or it can be saved to a file and run from the command-line for batch processing.

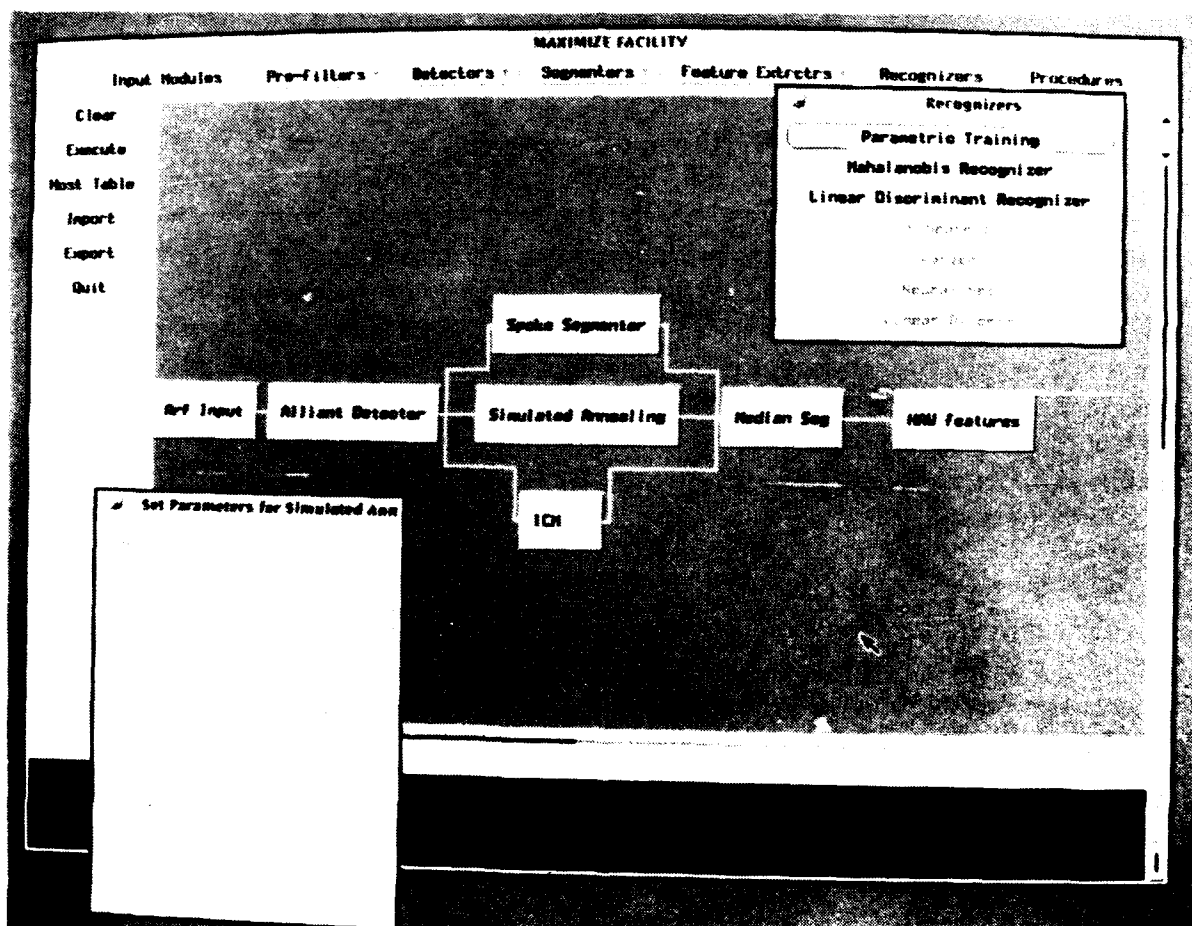


Figure 1: MAXIMIZE Graphic User-Interface

To get statistically significant results, most useful experiments in ATR research require processing thousands of images. If this were to be performed on a single workstation, it might take weeks or even months to perform a single experiment. In the past, special-purpose hardware has been developed for use in such experiments. While this reduces the amount of time of an experiment, the disadvantages are unportable software, longer development times and reduced flexibility in software development. Eventually obsolete hardware results in systems which are unreasonable for continued software development.

A network of algorithm components may have an inherent parallel architecture such as the ATR in Figure 1. In addition, most ATR experiments use a time-sequence of images which allows for pipelining of processing. The MAXIMIZE facility can recognize both of these opportunities for concurrency and will distribute the workload of an ATR across a network of general-purpose computers. This combines the power of a super-computer with the portability of a standard high-level programming language for ATR development and testing.

IV. Experiment Description

An experiment was conducted where three segmenters operated on the same input of interest points. The outputs of these three segmenters were then sent as input to a "Median Segmenter". The Median Segmenter uses the output of the three different segmentation techniques to create its own segment of an interest point. The criteria it uses for combining segments is simple: if two of the three segmenters identified a pixel as a target-pixel, then the Median Segmenter identifies that pixel as part of its segment. Historically segmenters usually identify all or most the pixels on a target; however, they often wrongly include parts of the background in the segment. The Median Segmenter works on the assumption that two independent segmenters will not both choose the *same* wrong pixels. The same features were then calculated for each of these four segmentation techniques and statistics were collected on these feature values.

These statistics were collected for over 10,000 images such as the one in Figure 2. These inverse-video images were taken from the CECOM Night Vision & Electro-Optics Directorate's Terrain Board for simulating infrared images. Targets in the images included M-1A1, M-60A3, M-2, M-113, M-35A, HUMMV, M-163, M-730-A1, A-10, and F-15. Since ATR performance is often affected by the number of non-target objects in an image, such as trees and rocks, the images were collected over three different levels of such "clutter". These levels of clutter are loosely defined as "high," "medium," and "low." All the targets in the images had a simulated range of about 2500 meters.

The mean and covariance matrix of the features were calculated for each of the targets at each of the clutter levels as well as for all the clutter levels combined. The Bhattacharyya distance¹ was used to measure separability of the target classes. The Bhattacharyya distance for a two-class problem is

$$\mu\left(\frac{1}{2}\right) = \frac{1}{8} (M_2 - M_1)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (M_2 - M_1) + \frac{1}{2} \ln \frac{\left| \frac{\Sigma_1 + \Sigma_2}{2} \right|}{|\Sigma_1|^{\frac{1}{2}} |\Sigma_2|^{\frac{1}{2}}}$$

where M_x and Σ_x are the mean feature vector and covariance matrix for a class.



Figure 2: Typical image from Night Vision & Electro-Optics Directorate's Terrain Board

For this experiment ten targets were used which results in 45 different two-class problems. The Bhattacharyya distance was calculated for each two-class problems. The average of the two-class Bhattacharyya distances for each of the segmentation procedures is shown in Figure 3. Figures 4, 5 and 6 show the averages of the two-class Bhattacharyya distances at low, medium, and high clutter levels, respectively.

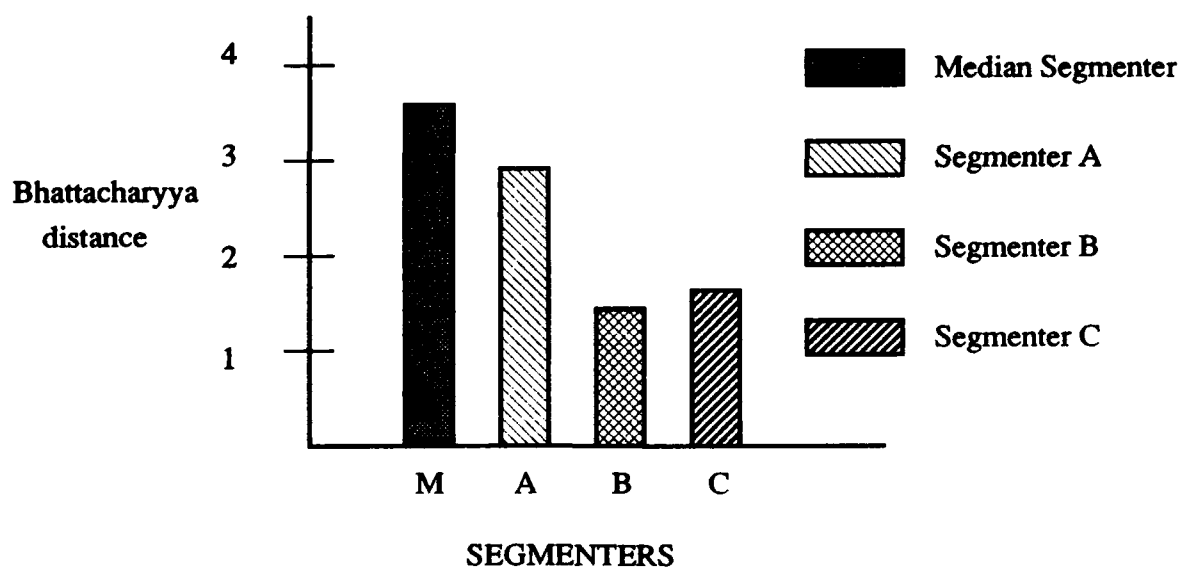


Figure 3: Average Bhattacharyya Distance for all clutter levels

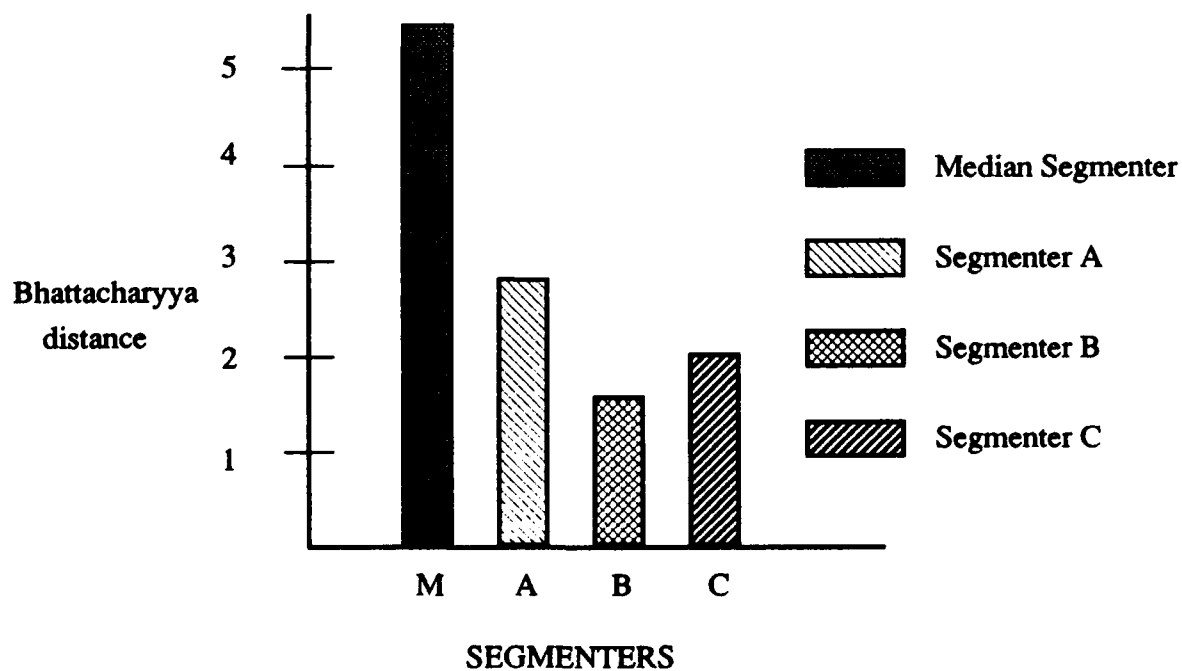


Figure 4: Average Bhattacharyya Distance for low clutter

DE CAMP

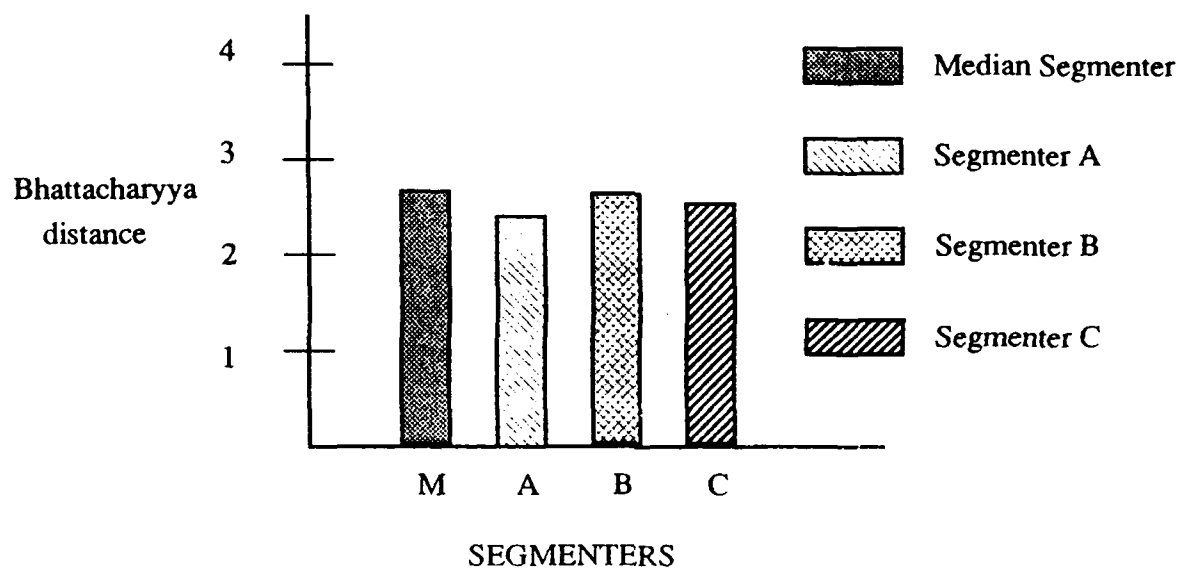


Figure 5: Average Bhattacharyya Distance for medium clutter

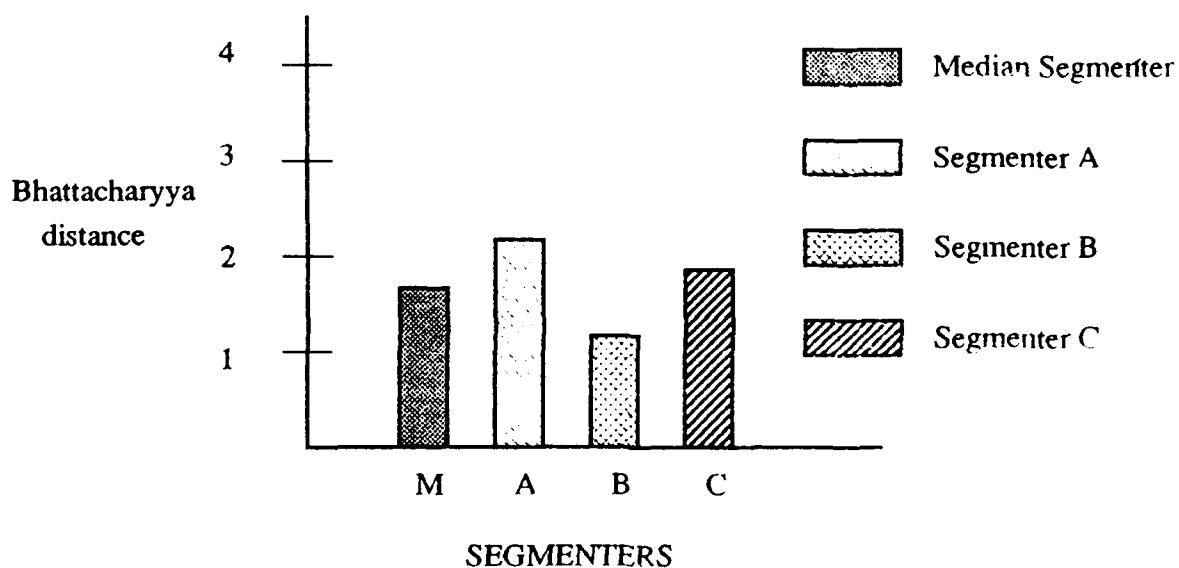


Figure 6: Average Bhattacharyya Distance for high clutter

V. Conclusions

The Median Segmenter provided the best overall separability between classes. This is largely a result of its superior performance in the low clutter case. Segmenter A provided the most robust performance at the different clutter levels. It is not fully understood why the Median Segmenter was less effective than other techniques in high clutter. This will be the subject of a subsequent study.

Conducting an experiment such as this without using the MAXIMIZE facility would have been very difficult. Simply getting three segmenters to work together could take several weeks of risky modifications to source code. Then processing 10,000 images on a single machine could have taken months on a single computer. MAXIMIZE automatically distributed the processing on several computers and results were computed in just a few days. The development of this facility has vastly improved the Army's ability to develop ATR systems which will identify threats in future battlefields.

1. K. Fukunaga, *Introduction to Statistical Pattern Recognition* (Harcourt Brace Jovanovich, 1972), pp. 70-72.

DILLON

A Framework for the Environmental Risk Assessment of
Contaminated Sediments at DOD Installations

Dr. Tom M. Dillon, REP
USAE Waterways Experiment Station (CEWES-ES-R)
Vicksburg, MS 39180-6199

INTRODUCTION

Over 17,000 sites at more than 1,800 Department of Defense (DOD) installations are under consideration for remedial action because of chemicals released into the environment (1). Nearly 70% of DOD sites are located on Army installations. Assessment and remediation activities are carried out under the Defense Environmental Restoration Program (DERP). DERP is analogous to the civilian Superfund program but is specific to active and formerly used DOD installations. The two programs have interacted historically and use the same or similar evaluative procedures. Partly because of this shared history, DERP was included as §211 when Superfund was last reauthorization (Superfund Amendments and Reauthorization Act [SARA] of 1986 [PL 99-499]).

Both DERP and Superfund use a process called risk assessment to evaluate the potential harm posed by chemicals in the environment. Historically, risk assessments have focused on the risks of contaminated groundwater to human health. Today, there is an increasing concern that other environmental media such as soils and sediments are highly contaminated and that nonhuman biological receptors may also be at considerable risk.

Many DOD installations have contaminated sediments on-post or adjacent to installation property. Concentrations of chemicals in sediments can be one to several orders of magnitude greater than those found in adjacent water bodies. These elevated concentrations are the result of physico-chemical factors such as organic carbon content, electrostatic attraction and hydrophobic repulsion. Consequently, sediments tend to act as sinks for chemicals released into the environment (2). Do these higher chemical concentrations pose significantly greater risks to human and non-human target receptors? At the present time, there are no risk assessment procedures available to Army base commanders and DOD environmental

DILLON

managers for evaluating contaminated sediments under the DERP or Superfund program. This paper presents the first such framework for assessing the environmental risk posed by contaminated sediments. Because the framework generally follows the generic human health risk assessment process currently used in the Superfund and DERP programs, that approach is briefly described below.

The Human Health Risk Assessment Process

The basis for most of today's human health risk assessments is the approach proposed by the National Academy of Sciences (NAS) in 1983 (3). The process focused on the probability of humans getting cancer as a result of some chemical exposure. The U.S. Environmental Protection Agency (EPA) adopted the NAS human health paradigm and currently uses it in many of its programs including Superfund (4). The approach consists of four elements: a) Hazard Assessment, b) Exposure Assessment, c) Effects Assessment and d) Risk Characterization. Each element is briefly described below.

Causality is established in the first step; Hazard Assessment. That is, can the chemical hazard cause cancer under any circumstance. If it cannot, there is no reason to assess risk. If it can, the risk assessment process is continued to completion. The next two elements, Exposure Assessment and Effects Assessment, are conducted in parallel. In Exposure Assessment, the magnitude and duration of chemical exposure is quantified. Here, a distinction is made between pathways and routes. An exposure pathway is the spatial/temporal fate of a chemical hazard from its source to the point of human contact. The manner in which that chemical hazard enters human beings at the point of contact is referred to as the route of exposure (e.g., ingestion, inhalation, dermal absorption). In Effects Assessment, the cancer potency of individual chemicals is established via dose-response data. These data are generated by monitoring the incidence of cancer in human surrogate species (mice, rats, etc.) exposed in the laboratory to a series of very high chemical concentrations. High concentrations are needed to increase the test's sensitivity because the incidence of cancer is typically very low. Results are then extrapolated to lower more environmentally realistic exposure concentrations and from the surrogate test species to humans. Information from Exposure and Effects Assessments are combined in the last element of risk assessment, Risk Characterization. Risks are usually expressed quantitatively as the number of excess cancer deaths per year (e.g., 10^{-6} or one additional death per million individuals). The uncertainties associated with these numerical estimates of risk are also summarized at the Risk Characterization stage.

DILLON

The NAS human health risk assessment model represents a logical approach for evaluating human health chemical risks. It enjoys wide acceptance and recognition in the scientific, engineering, and regulatory communities. This is an important programmatic consideration for Army base commanders and DOD environmental managers. On the other hand, the protocol, as currently practiced, has several major constraints *vis a vis* the assessment of contaminated sediments.

The NAS risk assessment paradigm emphasizes a single endpoint (cancer) in a single species (*Homo sapiens*). It was not designed to consider nonhuman target species of concern nor other endpoints. This is a serious limitation since many Army installations provide critical habitats to commercially/ecologically important animals as well as threatened and endangered species. For these organisms, gross morbidity such as tumors and lesions may not be particularly relevant. Impacts on survival, growth, and reproduction, especially as they affect population viability, are much more ecologically meaningful endpoints.

Another serious impediment to using traditional risk assessments for contaminated sediments is its chemical-by-chemical approach. Sediments are mixtures of chemicals embedded in a highly complex and heterogeneous mineralogical matrix. The chemical-by-chemical approach cannot account for contaminant interactions such as synergisms and antagonisms. Moreover, a chemical-by-chemical approach to assessing contaminated sediments will always be incomplete. Bulk sediment chemical inventories are rarely exhaustive and the analyst may not even know what to look for. A chemical or chemicals may be present in the sediment and exerting toxic effects but is never analyzed. The limitations of the chemical-by-chemical approach is especially critical for Army sites which contain military-unique compounds. These chemicals are exotic and may be difficult to analyze. Moreover, their toxicological significance, especially to nonhuman receptors, is often unknown. Some alternative to the chemical-by-chemical approach is required if the traditional human health risk assessment paradigm is followed for contaminated sediment evaluations.

A final constraint is the highly conservative nature of traditional human health risk assessments and the expression of risk in absolute terms. Installation commanders and DOD environmental managers should be aware that the numerical results of a cancer risk assessment are rarely accurate. That is, the incidence of cancer "predicted" by the risk assessment will not necessarily occur. This is because the process involves a series of highly conservative assumptions, extrapolations and calculations. This conservative approach is taken in the face of large uncertainties. Consequently, numerical results of a human health risk assessment are upper bound estimates. The true cancer risk actually lies somewhere between zero and the numerical estimate.

DILLON

A FRAMEWORK FOR THE ENVIRONMENTAL RISK ASSESSMENT OF CONTAMINATED SEDIMENTS

Environmental Risk Assessment (ERA) as used here is defined as "the process of evaluating the potential harm to biological systems arising from the intrinsic toxicity of sediment and the probability of exposure." Both human and nonhuman target species are included under "biological systems". The term "environmental" distinguishes this framework from; a) the NAS paradigm that deals exclusively with human health and b) EPA's on-going Ecological Risk Assessment effort. The latter program specifically excludes human health considerations but does include nonchemical perturbations such as habitat destruction, global warming, etc. The ERA framework proposed here consists of four elements analogous to the NAS paradigm.

ERA Element 1. INITIAL EVALUATION

Purpose. The first step is to clearly state why the ERA is being conducted and how the results will be used. For example, is this a baseline ERA or one associated with a specific remedial action? As will be seen later, selection of a particular technical approach or evaluative procedure will be driven by the purpose of the ERA. Consequently, time spent here fully developing the "whys" and "hows" will greatly improve the technical quality of the ERA and facilitate decision-making at its conclusion. Conversely, failure to fully develop the ERA's *raison d'être* will have negative repercussions throughout the process. One of the first steps in delineating the purpose is to identify assessment endpoints and measurement endpoints (5).

Assessment endpoints. Assessment endpoints are characteristics of the site or system that are highly valued and should be protected, restored or remediated. With regards to contaminated sediments, some valued characteristics of a site or system might be:

- (a) local fish and shellfish which are "safe" to eat
- (b) economically viable fish and shellfish industry
- (c) sustained populations of ecologically important species
- (d) "productive" benthic communities
- (e) bottom sediments which are not "toxic"
- (f) fish which are free of gross morbidity (e.g. lesions, tumors)

Value characteristics largely reflect societal concerns and drive the decision-making process. Consequently, the primary responsibility for developing appropriate assessment endpoints lies with the regulatory decision-maker. He or she is also charged with communicating the technical results to the public after the risk assessment is completed. By setting the assessment endpoints, the decision-maker helps to insure that the

DILLON

ERA is properly focused (see Risk Management below). In establishing assessment endpoints, the decision-maker should seek input from the technical community to insure assessment endpoints are biologically relevant, linked to the contaminated sediments and can be measured.

Measurement endpoints. Measurement endpoints are discrete observations which quantify changes in the value characteristics of the site or system. Ideally, measurement endpoints are the same as the assessment endpoints. This, however, almost never happens because assessment endpoints are either not measurable, or their measurement would be time- and resource-prohibitive. For example, it would be very difficult to measure all things necessary to causally link the effects of contaminated sediments to a diminution in a sport or commercial fishery. Measurement endpoints, therefore, are surrogates for assessment endpoints. In contrast with the latter, measurement endpoints are identified by the technical community. It is also their responsibility to inform the decision-maker if a particular assessment endpoint appears inappropriate. To that extent, development of assessment and measurement endpoints is an interactive process involving an active dialog between the regulatory and technical communities.

Receptor characterization. Receptor characterization is the identification and description of potentially exposed populations of concern. There are two broad categories of receptors, human and nonhuman. It is important to identify receptors of concern early in the ERA process because subsequent technical evaluations will differ significantly. For example, ERAs involving human receptors utilize a chemical-by-chemical approach and focus on a single endpoint; cancer.

Nonhuman receptor characterization may involve one or more target species depending on the goals of the ERA. For example, if the concern is for a particular commercial fishery, the receptor of concern is likely to be a single species. On the other hand, if the focus is maintaining a productive benthic community, a number of target species will be evaluated.

The accuracy of the ERA for contaminated sediments may be greater when nonhuman target species are involved. For example, the species of concern can be tested directly with no extrapolation. This is never an option for human health evaluations. Even if nonhuman surrogates are used, representative species are available (6). This is especially important when threatened or endangered species are involved. Effects can be determined under conditions which closely mimic *in-situ* exposures. This too, is never an option with human health risk assessments. The tedious chemical-by-chemical approach used for human receptors can often be abandoned in favor of an effects-based approach which is more appropriate for

DILLON

complex mixtures (see **EFFECTS ASSESSMENT** below). Finally, field verification of real-world exposures can be carried out with greater accuracy and precision than possible with human epidemiological studies.

Qualitative ERA. The purpose of the qualitative ERA is to list and describe the contents of each element prior to the quantitative ERA. One of the first lists prepared was just discussed; the receptors of concern and any appropriate surrogate test species. The second list contains the chemicals of concern in the sediment including conventional pollutants as well as any military-unique compounds such as munitions, smoke/incendiary materials, lethal and incapacitating agents and their breakdown products. If human receptors are involved, these chemicals are classified according to their known or suspected carcinogenic potential. For the quantitative **EXPOSURE ASSESSMENT**, the possible contaminant pathways and routes of exposure are listed and described for each target species of concern. The type of dose-response and/or exposure-response information needed to conduct the **EFFECTS ASSESSMENT** are described. Finally, known or suspected sources of uncertainty to be discussed during **RISK CHARACTERIZATION** are listed and described.

The qualitative ERA serves an important blueprint for the subsequent quantitative ERA. Once the latter begins, the **EXPOSURE ASSESSMENTS** and **EFFECTS ASSESSMENTS** tend to take independently, albeit parallel, paths. It is during the qualitative ERA that changes in either could be made to the benefit of both. For example, it may be possible to include an environmentally realistic exposure in the **EFFECTS ASSESSMENT** work of non-human target species. Also, the types of sediment exposures one might predominant could affect species selection for **EFFECTS ASSESSMENT**.

Quality assurance/quality control. Quality assurance (QA) is any formal program designed to generate reliable, technically sound data with accuracy and precision. Quality control (QC) refers to the specific steps taken to insure a successful QA program (7). An essential component in the QA/QC for Superfund sites are the Data Quality Objectives (DQO). The DQO are narrative and numerical criteria developed prior to data collection, which describe the qualitative and quantitative nature of the data required for decision-making during remedial response activities (8). It is beyond the scope of this paper to discuss the detailed DQO and QA/QC required for conducting ERA with contaminated sediments. Both issues, however, are addressed explicitly in the **INITIAL EVALUATION**.

Reference sediment. A reference sediment is used to interpret the significance of results obtained with contaminated sediments. This ERA, therefore, expresses the potential harm of contaminated sediments on an *incremental* or *comparative* risk basis. This is a significant departure from traditional human health risk assessments where risks are couched in absolute terms (e.g. 10^{-6} excess cancer deaths annually). By using a

DILLON

reference sediment and expressing results on an incremental of comparative risk basis, many of the highly conservative assumptions, estimates and calculations become moot. This is because they are common to the estimates of risk for both the reference and contaminated sediments. Solving simultaneous mathematical equations is a approximate analogy.

This comparative approach makes selection of the reference sediment absolutely critical to the ERA. Although responsibility for reference sediment selection lies with the technical community, the reference sediment should reflect the management goals and assessment endpoints developed by the decision-maker. For instance, management goals of "clean-up to ambient" versus "return to pristine conditions" require very different reference sediments. The reference sediment is defined as "a sediment which reflects the environmental conditions in the vicinity of the contaminated sediments in question". In many instances, the reference sediment is not contaminant-free.

ERA Element 2. EXPOSURE ASSESSMENT

EXPOSURE ASSESSMENT is defined as "the determination of spatial and temporal distributions of sediment-associated contaminants in relation to potential biological receptors of concern". As noted earlier, a distinction is made between pathways (chemical fate from source to receptor) and routes (method of entry into receptor). Elements of the **EXPOSURE ASSESSMENT** were listed and described previously during the qualitative ERA. There are two broad categories of **EXPOSURE ASSESSMENT** for contaminated sediments; direct and indirect. All **EXPOSURE ASSESSMENTS** for the ERA begin with an evaluation of direct sediment exposure.

Direct exposure. Direct exposure refers to the potential uptake of sediment-associated contaminants by nonhuman receptors inhabiting, ingesting or otherwise intimately associating with the sediments in question. The pathway and route are, for practical purposes, one in the same. Evaluating direct sediment exposure potential is a critical first step in all **EXPOSURE ASSESSMENTS**. If the chemicals of concern are not capable of moving from the sediment matrix into biota, then no exposure to any subsequent receptor is possible. Consequently, there is no risk. The lack of substantial bioaccumulation from contaminated sediments is not an uncommon occurrence since many chemicals are very tightly bound to the sediment matrix and not readily bioavailable (9).

Primary drivers in direct sediment exposure assessment are the spatial and temporal distributions of: 1) the target receptor and 2) the contaminated sediments. Receptors with the highest exposure potential in space and time are benthic infaunal organisms with limited or no pelagic life stage. For other organisms, exposure can be space- but not time-limited. An example would be an organism which spends most or all of its

DILLON

life in a spatially discrete habitat which encompasses only a small portion of the contaminated sediments in question. For some receptors, exposure is time-limited but not space-limited. An example would be an organism (e.g., a seasonal transient) which spends only a limited portion of its life associated with the contaminated sediments.

Indirect exposure. All indirect exposures involve the ingestion of biological material containing chemicals originating from the contaminated sediments in question. Not surprisingly, these assessments are more difficult to quantify and have much larger uncertainties. For nonhuman receptors, the pathways and routes are highly species- and site-specific. One must have knowledge of ingestion patterns of the receptor, its trophic relation in the local food web as well as chemical-specific trophic transfer potential. For example, assessment indirect exposure potential for a bottom scavenger such as a catfish are quite different from an anadromous top predator fish such as striped bass. As with direct sediment exposures, indirect exposures assessment consider both time and space distributions.

Indirect sediment exposure assessments for human receptors consider a specific route; the ingestion of contaminated fish and shellfish. Therefore, efforts are directed at identifying and quantifying the important pathways which precede that route. That information is often generated in the indirect exposure assessment if the nonhuman receptors is in the food web involving man (see preceding paragraph). Once a pathway to man is establish, human consumption patterns are determined. These include: 1) the proportion of fish and shellfish in the diet, 2) meal size and 3) frequency of meals. Potentially vulnerable sectors of the human population such as subsistence consumers, sport and commercial fishermen are given particular attention. This consumption information is combined with contaminant residue data for fish and shellfish to estimate a chemical-specific chronic daily intake (CDI). The CDI is used to characterize risks (see **RISK CHARACTERIZATION** below).

ERA Element 3. EFFECTS ASSESSMENT

EFFECTS ASSESSMENT is defined as "the process of establishing the strength or magnitude of the chemical hazard represented by contaminated sediments through the use of biological exposure-response or dose-response relationships". The specific technical approach depends on the target receptor of concern and the type of sediment exposure (direct or indirect). Appropriate test species were selected during the **INITIAL EVALUATION**.

Nonhuman target receptors. An effects-based bioassay approach is used for nonhuman target species exposed directly to contaminated sediment. The effects-based approach is not chemical-specific and considers the whole

DILLON

sediment as the potential hazard, rather than the individual chemicals. Effects-based sediment bioassays are preferred over the chemical-by-chemical approach for several reasons. One, many of the problems posed by chemical mixtures which cannot be overcome by the chemical-by-chemical procedure are circumvented. For example, sediment bioassays account for all chemicals (both known and unknown) which are present in the sediment and exerting an influence. Effects-based bioassays also account for all contaminant interactions (e.g. synergism). Two, sediment bioassays account for differences in bioavailability known to exist among various classes of contaminants. Three, there is no extant dose-response database analogous to human health which allows one to interpret the toxicological significance of specific chemical residues in nonhuman receptors. The time and cost required to generate such a database, even for a few animal-chemical combinations, would be prohibitive. These are important considerations for ERAs conducted at Army installations since sediments may contain a diverse and exotic mixture of chemicals (e.g., priority pollutants, munitions, smoke/incendiary materials, lethal and incapacitating agents and their breakdown products, etc.). Finally, the effects-based bioassay approach allows the risk assessor flexibility to examine a wide variety of different test species and endpoints of concern.

What "effects" are measured in these effects-based bioassays? Historically, survival has been measured in aquatic animals following acute (≤ 10 days) exposures. These acute lethality sediment bioassays, however, are limited to a very few, highly sensitive test species. Also, they do not simulate the *in-situ* contaminant exposures experienced by most benthic organisms; that is, the slow accumulation of contaminants over an extended period of time. Animals exposed to such conditions rarely die, but often respond in a more subtle sublethal manner. Consequently, a second generation of effects-based sediment bioassays is now being developed (10). These tests involve chronic sediment exposures and measure biologically important sublethal endpoints such as growth and reproduction. They also involve a more phylogenetically diverse group of organisms. PC-based population demographic models are being used to help interpret the results of these chronic sublethal sediment bioassays in terms of potential population-level impacts.

Sediment bioassays are used in exposure-response experiments to "establish the strength or magnitude of the chemical hazard represented by contaminated sediments". These experiments, which were initially designed during the qualitative ERA, are a function of the test species, its expected *in-situ* exposure and site-specific conditions. It is desirable, although not an absolute requirement, that the experimental design include exposure conditions one might expect in the field. The design must include, however, sediment exposures which will elicit some biological

DILLON

response. Otherwise, one will be unable to establish a relationship between exposure and response. Potential experimental designs may include one or more combinations of the following.

(a) Solid phase/single sediment - Test sediment is proportionally diluted with a suitable reference sediment.

(b) Solid phase/multiple sediments - Test sediments are ranked toxicologically and a suitable reference sediment is used to calibrate the ranking.

(c) Suspended phase/single concentration - A level of suspended sediment representing a particular field condition (e.g. storm event, typical near bottom conditions, etc.) is held constant and the proportion of test and reference sediment is varied.

(d) Suspended sediment/multiple concentrations - The proportion of test and reference sediment is held constant (usually 100%) and a gradient of suspended sediment concentrations is imposed.

Exposure-response experiments evaluating indirect sediment exposures and nonhuman target receptors involve contaminant trophic transfer studies. The specific design depends on the receptor of concern and its trophic position in the site-specific food web. For example, simple experiments involving just two trophic levels, use one of the designs in the previous paragraph to "contaminant" the food source. The exposure-response relationship is then evaluated by monitoring the response of the target receptor provided proportional amounts of this contaminated food source. Obviously, establishing the exposure-response relationship becomes more difficult as the number of trophic levels between the contaminated sediment and receptor of concern increase.

Human health. Human health **EFFECTS ASSESSMENT** for contaminated sediments is chemical-specific and does not differ substantially from EPA's recommended approach currently used at Superfund sites (11). The major difference is a much greater emphasis on the ingestion of fish and shellfish as the route of exposure. In fact, other routes (dermal, inhalation, ingestion of water or sediment) are ignored, unless there is a clear and ample justification for their inclusion. Contaminants of concern are classified as known or suspected carcinogens or noncarcinogens. Hazard strength is expressed as cancer potency factors and reference doses, respectively. Guidance regarding carcinogen classification, potency factors and reference doses is developed by EPA based on their review of laboratory dose-response experiments, epidemiological studies and other pertinent information. This guidance is periodically updated and made available through EPA's Integrated Risk Information System (IRIS) database.

DILLON

ERA Element 4. RISK CHARACTERIZATION

Risk Characterization is defined as "the technical coupling of **EFFECTS ASSESSMENT** and **EXPOSURE ASSESSMENT** to yield estimates of environmental risk for a particular receptor including a description of the uncertainties associated with these risks." In contrast to traditional human health risk assessments conducted at Superfund sites, risk is not expressed in absolute terms but as *incremental* or *comparative* risk. This is done by comparing risk results for contaminated sediments with similar results obtained with a reference sediment.

Nonhuman target receptors. Characterizing the risk of contaminated sediments to nonhuman receptors is dependent on the target species, its mode of sediment exposure (direct vs. indirect) and site-specific conditions. The exposure-response **EFFECTS ASSESSMENT** data generated in the laboratory are coupled quantitatively with the **EXPOSURE ASSESSMENT** field information. In some instances, conservative chemical tags are used to link laboratory exposures with field exposures. In some instances, narrative estimates of environmental risk are the only possible expression. This is particularly true when indirect sediment exposures and multiple trophic levels are involved. In all instances, risks to nonhuman target receptors are characterized spatially and temporally.

Human health. Estimates of chemical-specific risk are carried out generally as recommended by EPA (11). The primary difference is a major emphasis on the ingestion of contaminated fish and shellfish. The CDI is combined with toxicity information (slope factors or reference doses) to yield a chemical-specific risk.

Uncertainty analysis. By definition, all risk assessments are uncertain. Sources of uncertainty include: (a) *ignorance* of that which can be determined, (b) *error* in gathering knowledge, and (c) *stochasticity* or that natural variability which is unpredictable (12). The extent and magnitude of uncertainty is determined via uncertainty analysis (UA). These analyses can be qualitative, quantitative or a combination of both.

Qualitative UA is the most common and easiest to accomplish. In some instances, the qualitative UA provides the decision-maker with more usable information than a quantitative analysis. See Finkel (1990) for an excellent discussion of this subject (13). Some examples of qualitative UA include:

- (a) professional judgment
- (b) safety factors
- (c) listing of assumptions

DILLON

Quantitative UA requires much more data, a knowledge of the distribution of those data and relies heavily on numerical models. Consequently, they are conducted infrequently. Two important sources of uncertainty (other than stochasticity) are selection of the appropriate model and accuracy of the model parameters. A hierarchical list of quantitative UA are shown below.

- (a) statistical confidence levels
- (b) model substitution
- (c) sensitivity analysis
- (d) Monte Carlo simulations
- (e) error propagation analysis
- (f) field verification

RISK MANAGEMENT

Results of the ERA are used to help manage the environmental risks posed by the contaminated sediments. If the ERA indicates minimal or ephemeral risk, then no action may be required. If risks appear substantive or if critical receptors are involved, one or more remedial actions may be considered. The ERA can be repeated to consider exposure scenarios associated with different remedial actions. The ERA, however, is not the sole driver in risk management. It is collateral, or at times subordinate to other considerations such as costs or engineering feasibility. Risk management, therefore, can be defined as "the process of selecting the most appropriate regulatory action based on the ERA, applicable environmental laws and regulations, engineering feasibility, social values, economics, and other considerations". Risk management is the sole responsibility of the decision-maker.

RISK ANALYSIS

To complete the process represented by the ERA and RISK MANAGEMENT described above, one must field verify that (a) the estimates of environmental risks predicted by the ERA were accurate and (b) the actions taken by the decision-maker to manage those risks were effective. This verification can only be achieved through a well-designed field monitoring program based on the scientific method. Historically, monitoring programs have earned (rightly so) a terrible reputation. They expend large sums of money and accumulate large databases but are unable to provide definitive answers to decision-makers. To avoid this mistake, monitoring programs must be based on a series of *testable* hypotheses. Since avoidance of adverse impacts is highly desirable, the null hypothesis is the appropriate formulation. These hypotheses must reflect the assessment endpoints developed in the INITIAL EVALUATION. Levels which constitute acceptable

DILLON

or unacceptable change in the field are determined *a priori*, before the first data are collected. Field verification information provides the necessary positive feedback to modify and improve the ERA and RISK MANAGEMENT activities. This process is called RISK ANALYSIS.

CONCLUSIONS

1. A framework for the Environmental Risk Assessment (ERA) of contaminated sediments at DOD Installations is presented.
2. This ERA is driven by the concerns of the decision-maker. As such, heavy emphasis is placed on the initial development of assessment endpoints by the decision-maker and measurement endpoints by the technical community.
3. This ERA consists of four elements: INITIAL EVALUATION, EXPOSURE ASSESSMENT, EFFECTS ASSESSMENT and RISK CHARACTERIZATION. These elements are analogous to those found in traditional human health risk assessments conducted at Superfund sites.
4. This ERA, however, differs from traditional human health risk assessments in several respects.
 - (a) Both human and nonhuman target receptors of concern. This is important since many Army installations provide critical habitats to commercially or ecologically important animals, as well as threatened or endangered species.
 - (b) Test endpoints other than cancer are evaluated. Biologically important sublethal endpoints such as growth and reproduction are evaluated in a phylogenetically diverse suite of ecologically important species.
 - (c) Effects-based sediment bioassays are used to evaluate mixtures of chemicals embedded in a complex mineralogical matrix (sediments). This represents a significant improvement over the chemical-by-chemical approach.
 - (d) The environmental risk posed by contaminated sediments is expressed as *incremental* or *comparative* risk rather than in absolute terms. This is accomplished through the use of appropriate reference sediments.
5. The uncertainties associated with the estimates of environmental risk are explicitly addressed. Qualitative as well as quantitative means of expressing uncertainty are presented.

DILLON

6. This ERA can be used to establish technically sound clean-up levels for contaminated sediments at DOD installations as well as compare remedial alternatives.

7. A well-designed field monitoring program based on the scientific method provides field verification that (a) estimates of environmental risks predicted by the ERA were accurate and (b) risk management activities were effective.

Acknowledgments. Permission was granted by the U.S. Army Corps of Engineers to publish this information.

REFERENCES

1. MG P.J. Offringa, Keynote Address in *Proceedings from the 15th Annual Army Environmental R&D Symposium* (US Army Corps of Engineers Toxic and Hazardous Materials Agency Report No. CETHA-TS-CR-91076, 1991), pp. 1-14.
2. R.A. Baker, *Contaminants and Sediments. Volume 1. Fate and Transport, Case Studies, Modeling, Toxicity* (Ann Arbor Science, Ann Arbor, MI, 1980) 558 pp.
3. National Academy of Sciences, *Risk Assessment in the Federal Government: Managing the Process* (National Academy Press, Washington, D.C., 1983) 191 pp.
4. Environmental Protection Agency. 1987. The Risk Assessment Guidelines of 1986. EPA/600/8-87/045.
5. Environmental Protection Agency. 1989. Ecological Assessment Hazardous Waste Sites: A Field and Laboratory Reference. EPA/600/3-89/013.
6. L.W. Barnthouse, G.W. Suter II and S.M. Bartell. *Chemosphere* 17,1487 (1988).
7. G.M. Rand and S.R. Petrocelli. *Fundamentals of Aquatic Toxicology: Methods and Applications* (Hemisphere Publishing Corp., New York, 1985), 666 pp.
8. Environmental Protection Agency. 1987. Data Quality Objectives for Remedial Response Activities: Development Process. EPA/540/G-87/003.

DILLON

9. K.L. Dickson, A.W. Maki and W.A. Brungs. *Fate and Effects of Sediment-Bound Chemicals in Aquatic Systems*, (Pergamon Press, New York, 1987), 449 pp.
10. T.M. Dillon, in *Environmental Toxicology and Risk Assessment: 2nd Volume*, ASTM STP 1173, J.W. Gorsuch et al., Eds. (American Society for Testing Materials, Philadelphia, 1993) (in press).
11. Environmental Protection Agency. 1989. Risk Assessment Guidance for Superfund, Volume I. Human Health Evaluation Manual (Part A). EPA/540/1-89/002.
12. G.W. Suter, in *Acting Under Uncertainty: Multidisciplinary Conceptions*, G.M. von Furstenberg, ed. (Kluwer Academic Publishers, Boston, 1990), pp. 203-230.
13. A.M. Finkel, *Confronting Uncertainty in Risk Management. A Guide for Decision-makers*. (Center for Risk Management, Resources for the Future, Washington, D.C., 1990), 68 pp.

DILLON

TITLE: A Framework for the Environmental Risk Assessment of
Contaminated Sediments at DOD Installations
Dr. Tom M. Dillon, REP, USAE Waterways Experiment Station
(CEWES-ES-R), Vicksburg, MS 39180-6199

ABSTRACT:

Over 17,000 sites at more than 1,800 Department of Defense (DOD) installations are being considered for remedial action because of chemicals released into the environment. Many of these sites contain heavily contaminated sediments. A framework for the evaluation of environmental risks associated with these contaminated sediments is presented. It is driven by the needs of the decision-maker and thus, places heavy emphasis on the initial development of assessment and measurement endpoints. This framework is similar in structure to traditional human health risk assessments but with some notable exceptions. For example, both human and non-human target receptors are considered. This is significant because Army installations provide critical habitats to commercially and ecologically important animals as well as threatened and endangered species. Two, an effects-based sediment bioassay approach is used to evaluate the complex mixtures of contaminants normally found in sediments. This is crucial for DOD installations where a variety of conventional pollutants and military-unique compounds exist (munitions, lethal/incapacitating agents, smoke/incendiary material, etc.). The effects-based sediment bioassays represent a significant improvement over the more traditional chemical-by-chemical approach. Reference sediments are used to evaluate incremental or comparative environmental risks rather than absolute risks. Qualitative and quantitative methods to explicitly address the uncertainties associated with estimates of environmental risk are presented. Finally, this framework can be used to compare alternative remedial actions as well as establish technically sound clean-up levels for contaminated sediments at DOD installations.

BIOGRAPHY:

PRESENT ASSIGNMENT: Team Leader, CEWES-ES-R

PAST EXPERIENCE: Research Biologist, CEWES-ES-R

DEGREES HELD: PhD (Marine Science), MS (Zoology), BS (Biology)

DON, TRESSLER, RAPACKI

Ballistic Performance – Material Property Relationships of Glass Armor Versus Shaped Charge Jets

Mr. R.C. Don, Mrs. L. Tressler, and Mr. E.J. Rapacki, Jr.*

Armor Mechanics Branch, Terminal Ballistics Division
U.S. Army Ballistic Research Laboratory
Aberdeen Proving Ground, MD 21005-5066, USA

INTRODUCTION

The enhanced ballistic performance of glass armor versus shaped charge jets has been known for some time. Several theories have been put forth to explain the defeat mechanism involved, such as: closure (via rebound) of the glass surrounding the penetration cavity¹; dilatancy of the glass after comminution; and jet breakup by reflection of entrained material from the rear surface (backup) armor plate². Little work has been done, however, to determine how the physical and mechanical properties of a glass formulation relate to its performance as armor. Correlations of performance with material properties could enable armor engineers, working in concert with materials engineers, to tailor the glass formulation in order to fully exploit the capabilities of glass armor against shaped charge jets.

The present study relates the ballistic performance of glass armor to physical and relatively low strain rate dynamic material properties. The glasses chosen for evaluation represent a fairly broad range of mass density, elastic modulus and strength. Chemically, the formulations include: fluoro-, boro-, and alumino-silicates, soda-lime, fused silica and several high lead content formulations. Dynamic material properties were measured by drop tower testing, obtaining load versus time, as well as recovered comminuted particles. Confined finite glass targets with instrumented steel surrounds were ballistically tested using nominally quarter-scale shaped charge jets. Measured surface strains on the confinement allowed computation of the magnitude of its deformation during penetration. Residual armor steel penetration was used to compute the glass armor's mass and space effectiveness factors, which were correlated with the material properties and

DON, TRESSLER, RAPACKI

confinement structural response. A statistical analysis of the correlations was performed, indicating which properties were most significant in enhancing ballistic performance.

MATERIALS AND METHODS

Materials

Table 1 lists the glass formulations used by code designation (Corning Incorporated) with their chemical classification, mass density and Young's modulus. Specimens for drop tower testing were machined (ground) into flat-ended cylinders, 12.7 mm diameter \times 31.75 mm long. Ballistic test specimens were similarly fabricated, but were 38.1 mm diameter \times 76.2 mm long. With the exception of the 9658 *Macor* machineable glass-ceramic, all of the glass-ceramics were ceramed by the appropriate heat treatment after machining.

Table 1: Descriptions, Properties of Glasses Chosen for Dynamic and Ballistic Testing

Corning Code	Chemical Description	Density (g/cc)	Modulus (GPa)
7740	Borosilicate (<i>Pyrex</i>)	2.23	62.8
9741	Soda-Alumina Borosilicate	2.17	49.7
7940	Fused Silica	2.20	72.4
0330	Lithium Alumino-silicate (glass-ceram)	2.53	86.9
9606	Magnesium Alumino-silicate (glass-ceram)	2.60	120.0
1720	Aluminosilicate	2.52	87.6
7971	Titanium Silicate	2.21	67.6
0081	Soda Lime	2.47	71.0
9658	Fluorosilicate (<i>Macor</i> glass-ceram)	2.52	64.1
8463	High % Lead	6.22	51.7
RG	High % Lead Radiation Shield Glass	6.23	50.0
7597/8	Pb-Zn Borate	6.40	----

(Note: *Pyrex* and *Macor* are trademarks of Corning Incorporated)

Each glass cylinder for ballistic testing was ground to match the inside diameter of the confinement. The lateral confinement consisted of a close fitting, thick walled AISI 4140 steel tube, and 25.4 mm thick rolled homogeneous armor (RHA) steel plate was used for each end confinement. The steel tube was heat treated after machining to match the hardness of the RHA, nominally BHN 340. Epon 815 epoxy with Versamid V40/140 hardener was used to fill voids between the glass and the

DON, TRESSLER, RAPACKI

steel to achieve the best possible intimate contact between the sample and its confinement. Typical bond line thickness was 70-75 μm . After assembly, the steel tubes were machined to the length of the glass cylinders.

Strain gages were applied to the outer surface of the lateral confinement to measure the dynamic deformation of the steel confinement during jet penetration of the glass. Three rosettes were applied with epoxy, 19.1 mm from each end and midway along the length of the tube. The strain gages were 90° T-rosettes with 0.60 mm square grids (Micro-Measurements EP-08-062TT-120), one gage oriented parallel to the axis of the glass cylinder, to measure longitudinal strain, and the other oriented to measure transverse (hoop) strain. In designing the confinement, the surrounds were sized according to the anticipated pressures that could be generated within the glass during defeat of the jet, but removed from the glass/jet interaction zone. Franz and Lawrence report pressures on the order of 0.3 GPa from experiments with manganin stress gages imbedded in soda-lime glass targets³. It was desired to keep the inner wall of the surround from yielding in order to simplify the analysis of the recorded strain signals. A simple analysis indicated that a wall thickness of 19.1 mm would satisfy the requirement.

The end confinement plates were welded to an additional structural steel ring to contain the confined glass sample. Thus, the confinement tube was not laterally constrained. The structural ring also provided convenient instrumentation cable feed-through connections (see Figure 1: Cutaway View of Instrumented Target).

The characteristics of the shaped charge used for ballistic testing are listed in Table 2. The 20 mm long lead tip was intended to be removed by the front 25.4 mm RHA plate of the target, so that the glass would interact only with stretching jet. The reference penetration, 162.5 mm, was measured in stacked plates of 25.4mm thick RHA of BHN 340 and reference mass density, $\rho_{\text{ref}} = 7.85 \text{ g/cc}$.

Table 2: Sub-Scale Shaped Charge Characteristics

Liner Material: Copper Thickness: 0.66 mm Diameter: 36.9 mm Geometry: 42 degree cone	Tip Diameter: 2.5 mm Length: 20.0 mm Velocity: 8.3 mm/ μs
Jet Diameter: 1.4 mm Explosive: PBX N5	Breakup Time: 62 μs Penetration: 162.5 +/- 2.4 mm RHA

Drop Tower Testing

Material properties at moderate strain rates were obtained from drop tower testing with an instrumented tup. Samples were placed into a tightly sealed anvil chamber in order to recover the fractured glass. The drop mass was 225 kg, released from a height of 30.5 cm. Figure 2 is a representative load versus time plot as recorded by the drop tower instrumentation. This data allowed computation of the peak stress, time to fracture, and elastic strain energy absorption for each glass. The fractured samples were carefully recovered and sieve analyzed for particle size distribution.

Shaped Charge Experiments

Ballistic tests were conducted with targets and shaped charges as described above, using a vertical shot line to facilitate charge to target alignment. Three replications were performed to obtain statistically meaningful results. The charge was carefully positioned at the prescribed cone-base to target standoff distance, and so that the jet path would coincide with the axis of the glass cylinder. An exploding bridge wire detonator was used to initiate the explosive. Several additional 25.4 mm thick RHA plates were added below the target so that residual penetration was into effectively semi-infinite armor steel.

Strain gage instrumentation consisted of a constant voltage (nominally 18 VDC), pulsed excitation bridge circuit to increase the signal to noise ratio⁴. The strain signals were recorded on Nicolet 2090 digital storage oscilloscopes, one rosette per oscilloscope. An instrumentation trigger signal was obtained via an inductive pickup on the detonator signal line. A break screen mounted on the target front allowed determination of the time interval between the trigger signal and lead tip arrival, typically 24 μ s for the standoff used. Data reduction of the recorded signals was performed using software created with National Instruments' *LabWindows 2* package. The computer routine included a low-pass filter to minimize very high frequency (> 1 MHz) noise, and gage transverse sensitivity was accounted for.

RESULTS AND DISCUSSION

The reduced results of the drop tower tests, sieve analyses of the recovered drop tower test samples and measured peak hoop strain of the confinement are presented in Table 3.

DON, TRESSLER, RAPACKI

Table 3: Dynamic properties obtained from drop tower and ballistic tests

Glass Code	Time to Fracture (μ s)	Peak Stress (MPa)	Elastic Strain Energy (Nm)	Median Dia(μ m)	Confinement Peak Strain
7740	510	1358	59	240	
9741	510	1179	70	220	0.3%
7940	440	1317	49	205	
0330	330	1165	30	450	0.5%
9606	510	1882	57	1400	
1720	230	1200	32	1200	
7971	390	1282	48	205	
0081	210	834	13	690	0.4%
9658	200	648	14	4500	
8463	130	462	10	910	
(RG)	160	503	12	800	2.0%
7597/8	120	455	9	1200	1.0%

Referring again to Figure 2, the load release in the instrumented tup is an artifact of the instrumentation, representing the release wave from the far end of the steel anvil. The fracture time is the time from load application to the onset of load decrease after peak load has been achieved. From the known drop-weight velocity, tup displacement was calculated up to the fracture time. The elastic strain energy is the integral under the load-displacement curve. Sieve analysis of the recovered comminuted particles indicated similarly shaped distribution curves; the median particle diameter was used in subsequent regression analyses.

Figure 3 shows hoop strain versus time for confinements of representative high and low density materials. Comparison of the strain signals revealed that the confinement of the high density glasses had much higher peak hoop strain (approximately three times larger) than the peak strain for the lower density glasses. Measurement of the high density glasses' confinements indicated some yielding of the confinement walls. The quasi-steady-state strain, i.e. the mean level following the peak strain, showed good agreement with postmortem measurements of permanent set.

In all cases, residual penetration beyond the glass into armor steel was obtained. The plates were sectioned to determine the depth of penetration. The ballistic performance of the glass materials was calculated with respect to the

DON, TRESSLER, RAPACKI

reference armor steel penetration of the shaped charge jet, $P_{ref} = 162.5$ mm. The space and mass effectiveness factors, e_s and e_m , of the glass were computed from:

$$e_s = \frac{P_{ref} - P_{stl}}{t_g} \quad \text{and:} \quad e_m = e_s * \frac{\rho_{ref}}{\rho_g}$$

where t_g is the thickness of the glass and P_{stl} the total armor steel thickness penetrated (front plate plus residual penetration), and ρ_{ref} and ρ_g are the mass densities of the reference armor steel and the particular glass, respectively. The overall efficiency is: $q^2 = e_s * e_m$. Figure 4 shows q^2 by glass type. The high density leaded glasses exhibited poor performance in relation to the other types. Note that there appear to be three performance groupings, with no overlap between them. The best performance was achieved with 7740 Pyrex.

The armor performance measurands e_m , e_s , and q^2 were correlated with linear regression fits to each of the material properties. Due to the distinctly different (lower) performance of the high density glasses, these were not included in the regression analyses, except for the peak hoop strain data. Figure 5 shows the correlation coefficients grouped by material property. In order to ascertain the statistical significance of the correlations, a Student's t -distribution test was performed at the 97.5% confidence level. The t_o found in the t -distribution tables at $v = N - 2$ is easily transformed to r_o , (the minimum correlation coefficient, r , at which to expect a significant correlation) by⁵:

$$|r_o| = \frac{t_o}{\sqrt{N - 2 + t_o^2}}$$

The r_o 's are indicated in Figure 5 by a box within each group. Correlation coefficients greater than r_o for a particular property indicate significant influence of that property on ballistic performance.

CONCLUSIONS

Correlations of material properties of various glass formulations versus overall ballistic performance against shaped charge jets have been performed. The material properties were obtained using a drop tower impact test, and the ballistic tests were performed using sub-scale shaped charges against highly confined glass cylinders. Positive performance trends were observed for time to fracture and elastic strain energy absorption. Negative trends were evident for median post-

DON, TRESSLER, RAPACKI

fracture particle size as well as mass density. A negative trend, although less statistically significant, was also observed for peak dynamic deformation of the confinement. No significant correlation with performance was observed for the peak stress (as measured in the drop tower tests) or the elastic modulus of the glasses.

ACKNOWLEDGEMENTS

The authors would like to thank Mssrs. Gordon Dudder and Ronald Page of the Battelle Pacific Northwest Laboratories, who aided in the selection of materials and performed the drop tower testing and associated analyses, and Mr. Dominic Diberardo, who exercised extreme care in applying the strain gages used in this study.

REFERENCES

- ¹ G. E. Hauver, et al., "Penetration of Shaped-Charge Jets into Glass and Crystalline Quartz," Technical Report BRL-TR-3273, U.S. Army Ballistic Research Laboratory, 1991
- ² W. S. de Rosset and A. B. Merendino, "Jet Defeat Mechanisms in Brittle Armor," 1984 U.S. Army Science Conference Proceedings
- ³ R. E. Franz and W. Lawrence, "Stress Measurements in Glass During Shaped-Charge Jet Penetrations," Memorandum Report BRL-MR-3518, U.S. Army Ballistic Research Laboratory, 1986
- ⁴ E. J. Rapacki, Jr., "Instrumentation Techniques for Measuring Large, High Rate Strains with Foil Resistance Strain Gages," Technical Report ARBRL-TR-02573, U.S. Army Ballistic Research Laboratory, 1984
- ⁵ W. H. Press, et al., *Numerical Recipes*, Cambridge University Press, New York, 1986, p. 486

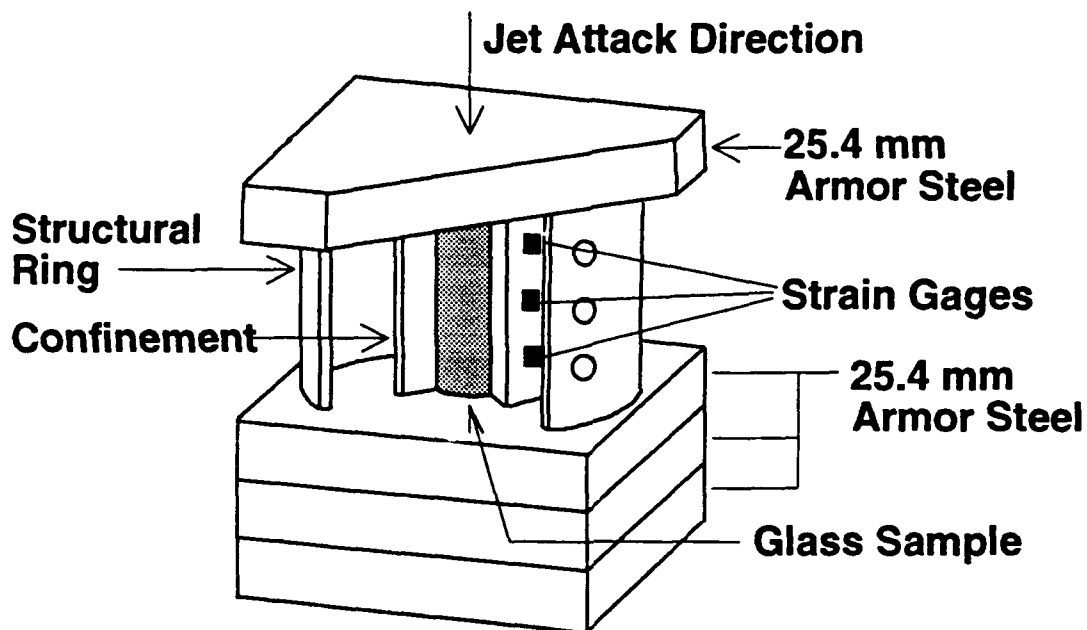


Figure 1: Cutaway drawing of the instrumented glass target

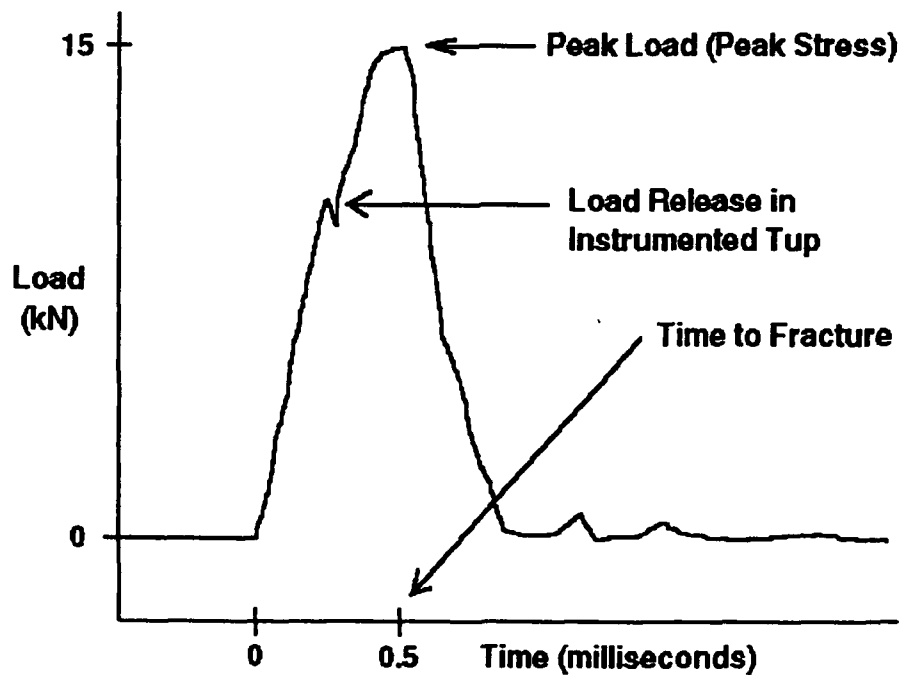


Figure 2: Representative curve of load vs. time from the instrumented drop tower tests

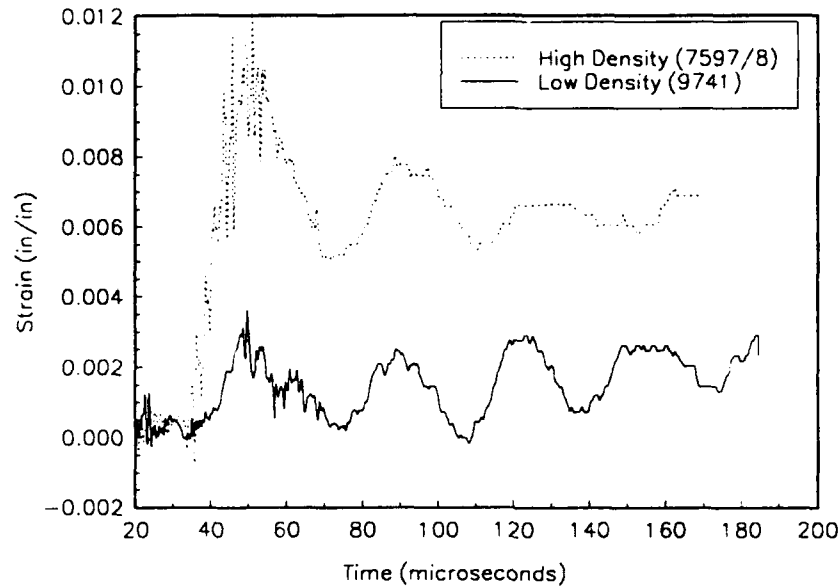


Figure 3: Confinement hoop strains for representative high and low density glasses

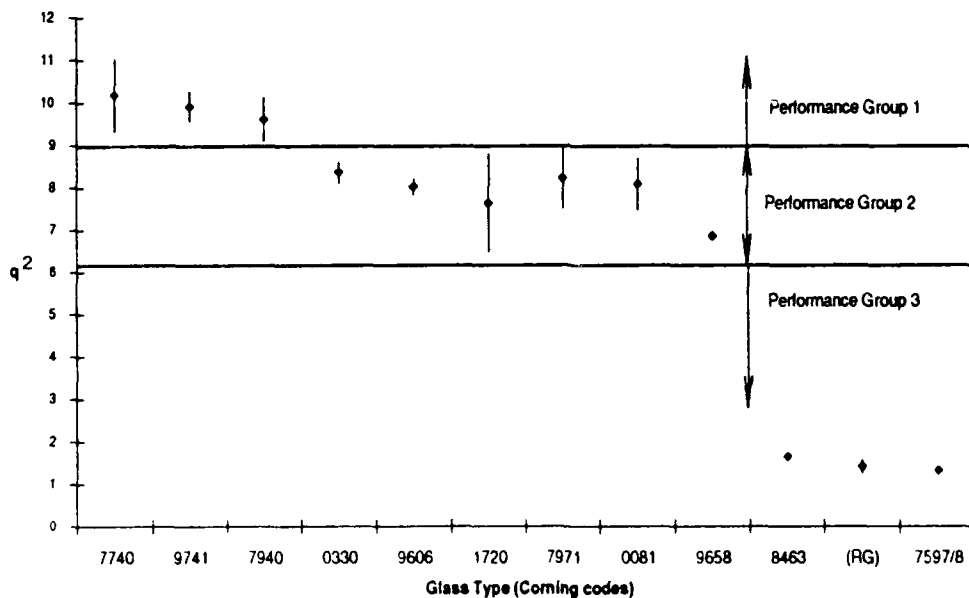


Figure 4: Glass armor ballistic efficiency, q^2 , by glass type (Corning Glass codes). Dots indicate the mean performance, while the attached bars indicate ± 1 standard deviation (N=3). Performance grouping is arbitrary (i.e. not done by any analytic means).

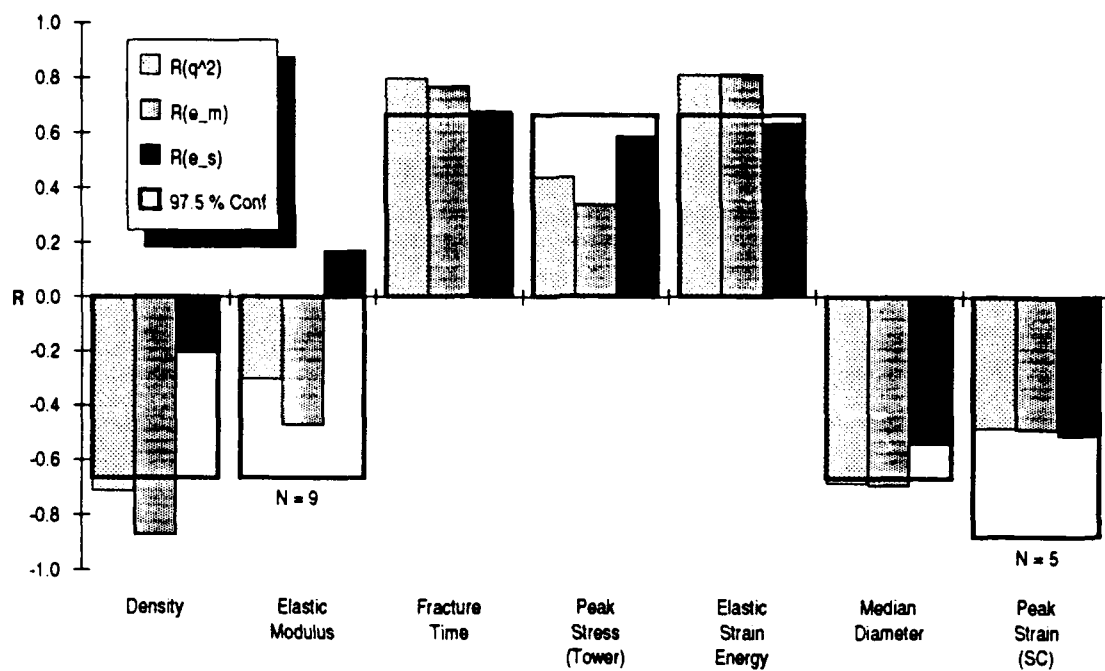


Figure 5: Correlation coefficients R for armor efficiency (q^2), mass effectiveness (e_m), and space effectiveness (e_s) indicated for each material property of interest.

DOWDING

Current Status of Tungsten Alloy Research

Mr. Robert J. Dowding
U.S. Army Materials Technology Laboratory
Watertown, MA 02172-0001

INTRODUCTION

The material selection criteria for kinetic energy (KE) penetrators requires that the chosen material possess high density, high strength, good toughness and adequate ductility. The material must also be corrosion resistant and readily available. There are several high density materials that are excluded due to their high cost. Examples are; gold, rhenium and platinum. The only metals, of high density, that meet these needs are depleted uranium (DU) and tungsten. These metals are alloyed to improve their mechanical properties while paying a small penalty in density. The current alloys of choice are uranium - 3/4 titanium (U - 3/4Ti) and tungsten heavy alloys (WHA). Most commonly, the WHAs are alloys of tungsten, nickel and iron, although copper, cobalt and silver are sometimes used. Typical mechanical properties for DU - 3/4 Ti and one commercial tungsten heavy alloy are shown in Table 1.

TABLE 1 PROPERTIES OF URANIUM AND TUNGSTEN ALLOYS		
	DU - 3/4 Ti	91 W-6.3 Ni-2.7 Fe
DENSITY	18.6 g/cc	17.2 g/cc
YIELD STRENGTH	900 MPa (130 KSI)	900 MPa (130 KSI)
TENSILE STRENGTH	965 MPa (140 KSI)	1100 MPa (160 KSI)
ELONGATION	24 %	12 %
HARDNESS	40 to 42 HRC	40 to 45 HRC

Quite naturally, both alloy classes have advantages and disadvantages associated with them. The tungsten alloys enjoy favorable characteristics related to health, safety and environment. Depleted uranium, a by-product of nuclear fuel enrichment, is an alpha particle emitter and consequently is slightly radioactive. Because of this, appropriate handling, storage and clean-up measures must be taken to prevent unmonitored release of the metal or its oxide. This includes the special steps that are taken in the transportation and disposal of waste DU. The result is added expense. But, the depleted uranium alloys are not without value, their greatest and most significant edge is in the area of armor penetration, where the DU alloys out-perform the tungsten heavy alloys. It has been demonstrated

DOWDING

that at ordnance velocities, the DU-based penetrators are able to penetrate up to 10% more armor steel (rolled homogenous armor) than tungsten-based penetrators. This performance gap is displayed in Figure 1 (Ref. 1).

The radioactive nature and associated handling, storage and disposal problems of the depleted uranium alloys leads to a strong desire to replace those alloys with ones of tungsten but, performance must not be sacrificed.

BACKGROUND

The search for tungsten alloys that ballistically perform as well as depleted uranium has, to date, been largely unsuccessful. The reason for this lies in the fact that until approximately 1990 it was believed that the performance gap could be reduced or eliminated by raising the level of the mechanical properties of the WHAs. A great deal of resources (financial and intellectual) were dedicated to understanding the processing and, the physical and mechanical metallurgy of the tungsten alloys. Through this effort significant gains were made in understanding the tungsten heavy alloy system. This work not only led to great improvements in the mechanical properties but also eliminated variability of the properties as a result of enhanced process control. Unfortunately, this effort did nothing substantial to improve the ballistic performance of tungsten materials. Table 2² lists the properties of several commercial tungsten heavy alloys that have been used for ballistic testing at The Army Materials Technology Laboratory (AMTL) over the past decade. The alloys shown represent the best commercially produced material available at the time it was purchased. Figure 2² displays the results of the ballistic testing of quarter-scale penetrators made from these heavy alloys. It is immediately obvious that the depth of penetration into semi-infinite rolled homogenous armor (RHA) is a function of velocity and density only. Similar data may be seen in Figure 3³.

Observation of the residual fragments from ballistic testing of tungsten alloy penetrators have been made by several researchers^{1,4-6}. Ekbohm, et al.⁴ state that the strength of a penetrator has little influence

TABLE 2
MECHANICAL PROPERTIES OF TUNGSTEN HEAVY ALLOYS
From Reference 2

GRADE	YIELD STRENGTH	TENSILE STRENGTH (KSI)	REDUCTION IN AREA (%)	ELONGATION (%)	IMPACT STRENGTH (FT-LB)
X21C	187.9	190.0	10.4	10.4	106
X27C	169.4	170.7	16.8	13.1	180
X27	171.0	174.0	20.4	-----	-----
W-2	149.2	150.3	1.8	3.1	-----

on the penetration properties into homogenous armor, an observation supported by others^{1,2}. Chaiat⁶ described the deformation observed after testing against the NATO heavy single target. He reports that all of the deformation occurs within 4 mm of the nose of the penetrator and the

DOWDING

microstructure of the remainder is as it was originally, showing no effect of the ballistic interaction. Chaiat's microstructural analysis further shows that within the first 4 mm the deformation occurs in bands rather than as uniform, general deformation. These observations, although quite valid, did not lead to greater understanding of the penetration process nor did they reveal fruitful areas of further research.

The real breakthrough occurred when side-by-side post-mortem analysis of tungsten alloy and depleted uranium alloy ballistic fragments allowed Magness and Farrand¹ to draw specific conclusions about the behavior of each of these alloys during penetration. They catalogued two types of fragments; those with mushroom noses and those with sharp noses. The sharp nosed fragments were characteristic of depleted uranium alloy penetrators while the tungsten alloy fragments had the mushroom noses. The mushroom formed at the tips of the tungsten alloy penetrators caused them to form penetration cavities that were larger in diameter than the cavities formed by the DU penetrators. Since conservation of energy considerations must apply, the cavities with smaller diameters were consequently deeper.

Examination of the microstructure of the ballistic fragments was particularly revealing in that the mechanism for maintaining the sharp nose on the DU penetrators was exposed. The mushroom petals sheared off before they could interfere with the penetration process and thus allowing a narrow diameter penetration cavity. The tungsten alloys, on the other hand, did not undergo this process a stable mushroom head was formed leading to a premature dissipation of the kinetic energy. It is believed that the shearing of the mushroom petals in the DU alloy is a result of adiabatic shear. This type of failure occurs at high strain rates and is described below.

ADIABATIC SHEAR

The deformation of a metal or alloy results in an increase in its strength through a process called work hardening. Work hardening is such an efficient process for increasing strength it is universally used. For example, a 90% tungsten heavy alloy in the as-sintered condition (no work hardening) has a yield strength of 85 ksi (600 MPa) and when worked to a 25% reduction in area its yield strength is raised to greater than 170 ksi (1170 MPa)⁷. In the deformation of any metal or alloy, the energy of deformation is usually dissipated in the form of heat. The thermal diffusivity of metals is higher than for most materials and when a metallic piece is deformed slowly (ie. a low strain rate) the heat is easily removed from the deformation zone. Deformation that occurs at elevated strain rates also generates heat but it can not be extracted quickly enough to prevent thermal softening⁸. Heating of this type is called, adiabatic⁹. Heating a metal usually leads to a reduction in the strength and in general, the higher the temperature the lower the strength¹⁰. A mechanical perturbation in the thermally softened zone causes the deformation to concentrate into a narrow region, or band, which leads to further heat generation. The snowball effect of deformation and heating results in highly deformed bands in the metallic piece and are called adiabatic shear bands. These shear bands can lead to unexpected failure and have been the subject of substantial research, predominately in the evaluation of ordnance and armor materials^{9,11-14}. A phenomenological description of the adiabatic shear instabilities that can occur in metals is shown in Equation One, where τ = shear stress, Γ = shear strain, $\dot{\Gamma}$ = shear strain rate and T = temperature.

DOWDING

$$d\tau = \left. \frac{\partial \tau}{\partial \Gamma} \right|_{\Gamma, \dot{\Gamma}} d\Gamma + \left. \frac{\partial \tau}{\partial \dot{\Gamma}} \right|_{\Gamma, \dot{\Gamma}} d\dot{\Gamma} + \left. \frac{\partial \tau}{\partial T} \right|_{\Gamma, \dot{\Gamma}} dT = 0 \quad \text{Equation 1}$$

The equation states that shear instability will occur when the sum of the terms is zero. In the case of these penetrator materials the strain hardening and strain rate hardening terms (the first and second terms) need to be balanced by the thermal softening (the third term) for instability.

TUNGSTEN HEAVY ALLOYS

Tungsten heavy alloys, as previously described, are the product of a powder metallurgy (P/M) manufacturing process. Elemental metal powders in the desired proportions are blended together, pressed to shaped and liquid phase sintered. During the sintering of these powders the tungsten particles grow from an initial 1 to 5 μm to greater than 40 μm in diameter. The procedure and process has been previously described in great detail and the interested reader is referred to recent articles on the subject¹⁵⁻¹⁶.

The microstructure of a typical tungsten heavy alloy is shown in Figure 4. It is a two-phase metal matrix composite formed *in situ* through alloying. The matrix (usually nickel, iron and all the tungsten it can hold in solution) is reinforced by the generally spherical tungsten grains. Unlike traditional metal matrix composites, the tungsten reinforcing phase of the heavy alloy makes up the greatest fraction of the volume and hence can dominate the properties.

DEVELOPMENT OF NEW ALLOYS AND COMPOSITES

Adiabatic shear is not observed in tungsten heavy alloys due to a combination of factors. The physical and mechanical properties as well as the state of stress play significant roles. In this application, there is little that can be done about the state of stress but the selection of a material that has more favorable properties can be made. There are two general approaches being taken in the effort to develop tungsten-based materials to replace the depleted uranium alloys used for penetrator applications. The first, attempts to identify an alternate matrix composition to substitute for the nickel-iron matrix while the second seeks to build an engineered structure that has selected failure zones.

While the nickel-iron matrix eventually shear localizes, the total strain required means that it does not do so at an advantageous point in the penetration event, it occurs much too late. Figure 5 shows an example of the deformation observed in a tungsten penetrator fragment. It is seen that locally intense deformation has occurred while other areas show little or no effect of the interaction. The matrix replacement notions are based on the suggestion that other materials may shear localize in a manner similar to depleted uranium. The only question appears to be whether or not they will shear localize in the presence of a significant volume of reinforcing phase (tungsten) and at the right time. The easiest method of demonstrating this concept is by making a tungsten composite in which the matrix is depleted uranium; the material which is known to exhibit the desired behavior under ballistic conditions. Work at Los Alamos National Laboratory (LANL) has accomplished just that. A sintered tungsten compact (72% dense) melt infiltrated with depleted uranium (28% by volume) was shown to perform as well as monolithic U - 3/4 Ti in a ballistic test. A flash radiograph of this demonstration penetrator after exiting the last

DOWDING

plate of the armor target showed the presence of a sharp nose, an indication of adiabatic shear (Private Communication, Mr. Bill Hogan, LANL, 1991). It is this proof of principle that is driving the search for alternate matrices for tungsten alloys and composites.

There are several material properties which govern the generation of adiabatic shear bands. They include; a low work hardening rate, low strain rate sensitivity and a high rate of thermal softening. These properties are also represented in the three terms of Equation 1 respectively. Other properties include microstructural stability and thermal properties such as thermal conductivity and heat capacity. The selection of candidate materials for alternate matrix compositions takes these factors into account. Other candidate materials are selected based on experience. For example, there is a large knowledge base in armor materials and how they behave under ballistic conditions and adiabatic shear has been observed in many of these materials. Unfortunately, there is no obvious prime contender to consider for use in penetrator applications. Table 3 lists several of the materials considered for substitution. For some of these combinations the systems can no longer be considered heavy alloys and will be referred to as composites.

TABLE 3 CANDIDATE MATERIALS FOR MATRIX SUBSTITUTION	
Metal or Alloy	Reason for Selection
Hafnium (Hf)	Properties Similar to Uranium
Titanium (Ti) Alloys	Exhibited Adiabatic Shear in Armor Applications
Copper (Cu) Alloys	Previous Observation of Adiabatic Shear
High Strength Steels	Exhibited Adiabatic Shear in Armor Applications
Intermetallics	Theoretical Treatment of Equation 1

There are many processing and alloying considerations that need to be resolved if the matrix replacement approach is to be successful. One of these is the grain size of the tungsten grains in the alloy/composite. In the tungsten heavy alloy these grains are from 30 to 50 μm in diameter and a typical adiabatic shear band in non-tungsten materials is often thinner than 10 μm . The question that arises is, does the grain size, of the tungsten alloy, need to be smaller than the expected width of the shear band and further, will the grain size have a limiting effect on the width of the shear band? There are some initial indications that the grain size may limit the width of the shear band. In torsional testing of a tungsten-pickel-iron (93% W) heavy alloy, at strain rates from 1×10^{-4} sec.⁻¹ to 100 sec.⁻¹, the width of the deformation band appeared to be approaching an asymptotic value of 50 to 75 μm ¹⁷. In other testing, a plug of tungsten

DOWDING

heavy alloy was punched from a thin disc (1/16 inch) at a strain rate of $1 \times 10^6 \text{ sec.}^{-1}$. The width of the shear deformation zone was approximately 40 to 60 μm ¹⁸. In both cases this value appears to be approaching the tungsten grain size. Some limited compression testing at a nominal strain rate of $5 \times 10^3 \text{ sec.}^{-1}$ indicated the importance of prior straining (cold working) on the intensity and width of the deformation band. The alloy tested was a tungsten-nickel-cobalt alloy and the prior working levels were 0, 7 and 25 percent reduction in area. The band widths were approximately 120, 50 and 25 μm wide respectively and it should be noted that the narrowest band had a crack at its centerline indicating very great intensity of deformation (Private Communication, Dr. Animesh Bose, Southwest Research Institute, San Antonio, TX, 1992). This limited work seems to indicate in order to get thinner, more intense, deformation bands a finer grain size and prior cold work is needed and that the grain size is dictated by the desired (or required) width of the shear band.

An examination of the tungsten grains in Figure 4 shows that many of them are in contact with their neighbors. This contiguity has been shown to be a weak link in the microstructure at low strain rates and its role at elevated rates of deformation is uncertain. The tungsten grain contact observed here forms a continuous network in the microstructure and may act to impede the propagation of any shear bands that form. If this microstructure is undesirable the elimination it is possible through the use of tungsten powder that has been coated with the desired matrix alloy. Development of this powder has been the subject of several efforts¹⁹⁻²⁵ and the microstructure of a tungsten heavy alloy produced by this method is shown in Figure 6. Note the complete absence of tungsten grain contact. Three point bend properties of this material also indicate that it is 50% stronger and has a greater deflection to failure than a traditionally produced heavy alloy like that shown in Figure 4. It can be envisioned that a microstructure of this type, no matter the composition of the matrix, would be beneficial to the propagation of a shear deformation band since it would provide a clear path in which interference from tungsten particles would be minimal.

STATUS OF CURRENT EFFORTS

Several possible matrix replacement materials can be identified if thermal diffusivity is examined. The thermal diffusivity of a material is a measure of how quickly heat is dispersed, reported here in m^2/sec . Equation 2, below, shows how the thermal diffusivity is calculated. This simplistic approach assumes that if the heat is not removed quickly enough then there is a chance for an adiabatic shear band to form. This approach also makes the assumption that the thermal diffusivity is the only factor controlling adiabatic shear band formation. As shown above, there are several factors to be concerned about in the generation of adiabatic shear bands. Examination of the thermal diffusivity helps identify candidate materials for consideration. Table 4 lists the thermal diffusivities for several metals and alloys that could be considered as matrix replacements. Also shown are tungsten, uranium and 70/30 nickel/iron alloy for baseline reference. The 70/30 nickel/iron alloy is the initial (prior to liquid phase sintering) composition of the heavy alloy matrix but, it must be remembered that this matrix after sintering contains all the tungsten it can hold in solid solution, about 25 percent by weight, which will certainly affect the thermal diffusivity. Where α = thermal diffusivity, k = thermal conductivity, ρ = density and c_p = specific heat of the material of interest.

$$\alpha = \frac{k}{\rho c_p}$$

Equation 2

TABLE 4
THERMAL DIFFUSIVITIES OF SELECTED METALS AND ALLOYS
From References 26 & 27

Metal or Alloy	U	W	Hf	Ti	Brass 70/30	Titanium Alloy (6Al-4V)	70/30 Ni/Fe
Thermal Diffusivity (m ² /sec x 10 ⁶)	13.0	67	12.3	9.1	34	5.8	6.6

Tungsten-Hafnium Alloys

From Table 4 it is seen that the thermal diffusivity of hafnium is slightly lower than that of uranium, making hafnium a candidate for matrix replacement. Monolithic hafnium has been shown to undergo adiabatic shear in high strain rate testing. Figure 7 shows a typical shear band in this material²⁸. Attempts at liquid phase sintering hafnium-tungsten alloys resulted in an embrittled alloy which was due to a high residual oxygen content. It was also reported that the tungsten phase acted as the matrix while the hafnium acted as the reinforcement which also contributed to the brittle behavior (Private Communication, Dr. Animesh Bose, Southwest Research Institute, San Antonio, TX, 1991).

The powder coating techniques reported above may be helpful in processing these alloys. This process along with a solid state, pressure assisted consolidation technique will eliminate all of the tungsten grain contacts. This will force the hafnium to act as the matrix. A secondary benefit from the solid state consolidation will be the retention of a fine grain size. This will promote higher strength and perhaps narrower, more intense shear bands. Preliminary work with this system has identified some of the chemical vapor deposition (CVD) parameters but has so far failed to deposit heavy coatings²². Lack of a thick coating has prevented determination of the shear localization tendencies of this system.

Tungsten-Titanium Alloys

Table 4 indicates that the thermal diffusivities of titanium and the titanium Ti-6Al-4V alloy have values for which shear localization might be expected. Adiabatic shear has been observed in titanium alloys⁹, particularly when used as armor. This prior observation of adiabatic shear due to ballistic interaction is particularly encouraging and it certainly indicates that shear localization will likely occur at the loading rates of interest. The greatest problem with the titanium alloys is their low density (4.5 g/cc). By comparison, the density of 70/30 nickel/iron is

DOWDING

8.59 g/cc and the density of hafnium is 13.1 g/cc. The low density issue will make the titanium materials a low priority for this application.

70/30 Brass

Copper and its alloys, through previous shear band observation²⁸, have also been identified as possible matrix substitutes and some brass alloys have been reported to show adiabatic shear bands (Private Communication, Dr. Charles White, Army Materials Technology Laboratory, Watertown, MA, 1990). A common brass alloy, that is available as a pre-alloyed powder, is 70/30 brass (70% copper, 30% zinc, by weight). By using pre-alloyed powder it was hoped that problems of zinc evaporation could be avoided since the melting point of zinc is quite low at 420°C. Sintering of tungsten-copper alloys is often performed in the range 1250°C to 1400°C²⁹. The sintering of tungsten-70/30 brass (with a 1% nickel addition to activate the sintering process), was attempted at temperatures ranging from 975°C to 1125°C, a further step to try to prevent vaporization of the zinc. This was not successful as the sintered densities were less than 90% of the theoretical densities and Table 5 shows the variation of zinc content with sintering temperature. In this table brass additions of either 10% or 20%, by weight, are shown which results in nominal zinc contents of 3% or 6%, by weight.

TABLE 5 ZINC CONTENT OF SINTERED TUNGSTEN-BRASS ALLOYS (By Weight)					
	Nominal Zinc Content (%)	975°C	1025°C	1075°C	1125°C
20% Brass	6	0.50%	0.49%	0.45%	0.01%
10% Brass	3	0.12%	0.02%	<0.01%	<0.01%

It is obvious, from these results, that for successful consolidation of tungsten-brass alloys that an encapsulated consolidation technique will be required. It's entirely likely that hot isostatic pressing (HIP) would be successful in obtaining full density while retaining the proper chemical composition but this will not be attempted immediately. It is believed that to get a cost competitive penetrator alloy the relatively simple technique of pressureless sintering must be used. Alternative copper-based matrix compositions are being considered and copper-niobium could offer a solution since it would allow much higher processing temperatures and hopefully eliminate the vaporization of critical alloying elements.

Intermetallic Compounds

The three components of Equation 1 describe the conditions necessary to observe adiabatic shear localization in a metal or alloy. For typical metals, the strength increases as the strain or strain rate increases and the strength decreases as the temperature increases. This is strain hardening, strain rate hardening and thermal softening respectively. Some materials, on the other hand, exhibit anomalous thermal strengthening, that is they get stronger as the temperature increases. Some of the intermetallic compounds display this behavior. Naturally, this increase in strength

DOWDING

is not without bound, at some temperature the ordered structure of the compound breaks down and rapid loss of strength is observed. As with typical metals and alloys, an increase in the strength is accompanied by a decrease in the ductility of the material. If the mechanism of shear localization is considered, where adiabatic heat is generated in a local deformation zone, it is seen that with these anomalous strengthening compounds the loss of ductility in a local region will accompany any increase in temperature and could lead to a cracking failure. Alternatively, if the structure does not crack, the temperature could increase to the point where the ordered structure is lost and sudden softening occurs leading to a local deformation band. Intermetallic compounds, specifically nickel aluminides, may be useful as matrix replacements since these compounds exhibit the described properties.

Intermetallic compounds will present unique processing challenges if they are to be successful matrix replacement materials. Fine aluminide powders are available³⁰ that could be used to sinter with tungsten powder but attaining full theoretical density could be difficult. The process of sintering relies on diffusion of adjoining powder particles with the ultimate goal of elimination of all porosity. The diffusivity of ordered materials is so low that full sintering will not occur in a pressureless environment. A pressure assisted technique will be required.

SUMMARY

It is highly desired to identify and develop the processing of tungsten-based alloys or composites that will place those materials in the technological position to replace depleted uranium (DU) in kinetic energy penetrator applications. The most viable options appear to be in the area of matrix replacement where the nickel-iron matrix is replaced with another material that mimics the favorable attributes of DU. These materials are hafnium, titanium and titanium alloys, copper alloys and intermetallic compounds.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the efforts of researchers at various contractors including; Drs. Jack Stiglich and Robert Tuffias and Mr. Brian Williams at Ultramet, Inc., Mr. Murray Kornhauser of 3C Systems, Drs. James Withers and Sumit Guha at MER Corporation, and Drs. Animesh Bose, James Lankford and Hervé Couque at Southwest Research Institute. Also acknowledged is the input and technical discussion of Dr. Tusit Weerasooriya, Messrs. Ernest Chin, Patrick Woolsey and Mark Sonnenfeld and especially Dr. Kenneth Tauer; all of AMTL. The author also acknowledges the support afforded by Messrs. John Nunes and Charles F. Hickey, Jr. of AMTL's Materials Producibility Branch. Lastly, the author notes the passing of Mr. Stanley L. Lopata (October 1991), he is and will continue to be missed.

DOWDING

REFERENCES

1. L.S. Magness and T.G. Farrand, *Proceedings of The 1990 Army Science Conference*, pp 465-479 (1991).
2. R.J. Dowding, K.J. Tauer, P. Woolsey and F.S. Hodi, *The Metallurgical and Ballistic Characterization of Quarter-Scale Tungsten Alloy Penetrators*, Army Materials Technology Laboratory, Watertown, MA, MTL TR 90-31 (1990).
3. C.E. Anderson, Jr., B.L. Morris and D.L. Littlefield, *A Penetration Mechanics Database*, Southwest Research Institute, San Antonio, TX, SwRI Report 3593/001 (1992).
4. L. Ekblom, S. Bogegård, L. Holmberg and L. Westerling, *The Proceedings of the 9th International Symposium on Ballistics*, pp 2-447 to 2-456 (1986).
5. U. Gerlach, *Met Trans A*, 17A, pp 435-442 (1986).
6. D. Chaiat, *Proceedings of the 1986 P/M in Defense Technology Seminar*, MPIF, (1986).
7. J.R. Spencer and J.A. Mullendore, *Relationship Between Composition, Structure, Properties, Thermo-Mechanical Processing and Ballistic Performance of Tungsten Heavy Alloys*, Contract DAAL04-86-C-0023, Army Materials Technology Laboratory, Watertown, MA, MTL TR 91-44 (1991).
8. J.F. Mescall and H. Rogers, *Role of Shear Instability in Ballistic Penetration*, Army Materials Technology Laboratory, Watertown, MA, MTL TR 89-104 (1989).
9. R. Dormeal, in *Materials at High Strain Rates*, pp 47-70, Elsevier Applied Science, London and New York (1987).
10. T.H. Courtney, *Mechanical Behavior of Materials*, McGraw-Hill Publishing Company, New York, (1990).
11. M.R. Staker, *Acta Metall*, 29, (1981), p 663.
12. M.R. Staker, *Scripta Metall*, 18, (1984), pp 735-738.
13. S.L. Semiatin, M.R. Staker and J.J. Jonas, *Acta Metall*, 32, (1984), pp 1347-1354.
14. C.L. Wittman, M.A. Meyers and H.-r. Pak, *Met. Trans. A*, 21A, (1990), pp 767.
15. R.J. Dowding, *P/M in Aerospace and Defense Technologies*, MPIF, Princeton, NJ, (1991).
16. L. Ekblom, *Scand J. Metall*, 20, (1991), pp 190-197.
17. T. Weerasooriya, P.A. Beaulieu and R. Swanson, *Effect of Strain Rate on Deformation and Failure Behavior of 93W-5Ni-2Fe at Room Temperature*, Army Materials Technology Laboratory, Watertown, MA, MTL TR in press.

DOWDING

18. M. Kornhauser and R.J. Dowding, *Development of Tungsten Based Composites*, Contract DAAL04-91-C-0012, Army Materials Technology Laboratory, Watertown, MA, MTL TR 92-7 (1992).
19. B.E. Williams, J.J. Stiglich and R.B. Kaplan, *Tungsten and Tungsten Alloys-Recent Advances*, A. Crowson and E. Chen, Eds., TMS, Warrendale, PA, (1991).
20. B.E. Williams, J.J. Stiglich, R.B. Kaplan and R.J. Dowding, *Tungsten and Tungsten Alloys-Recent Advances*, A. Crowson and E. Chen, Eds., TMS, Warrendale, PA, (1991).
21. B.E. Williams, J.J. Stiglich and R.B. Kaplan, *Coated Tungsten Powders for Advanced Ordnance Applications, Phase II*, Contract DAAL04-88-C-0030, Army Materials Technology Laboratory, Watertown, MA, MTL TR in press.
22. B.E. Williams and J.J. Stiglich, *Hafnium- and Titanium-Coated Tungsten Powders for Kinetic Energy Penetrators*, Contract DAAL04-91-C-0022, Army Materials Technology Laboratory, Watertown, MA, MTL TR in press.
23. S. Mulligan and R.J. Dowding, *Sinterability of Tungsten Powder CVD Coated with Nickel and Iron*, Army Materials Technology Laboratory, Watertown, MA, MTL TR 90-56 (1990).
24. B.E. Williams, J.J. Stiglich, R.B. Kaplan and R.H. Tuffias, *A Major Advance in Powder Metallurgy*, Presented at 2nd National Technology Transfer Conference (Technology 2001), (1991).
25. R.J. Dowding, *Tungsten and Tungsten Alloys-Recent Advances*, A. Crowson and E. Chen, Eds., TMS, Warrendale, PA, (1991).
26. F. Kreith and W.Z. Blacker, *Basic Heat Transfer*, Harper and Row, Publishers, New York, (1980).
27. C.J. Smithells, *Metals Reference Handbook, 5th Ed.*, Butterworths, London, (1976).
28. M.E. Blackman and S.E. Finnegan, in *Metallurgical Effects at High Strain Rates*, R.W. Rhode, et al. Eds., Plenum Press, New York, (1973).
29. J.L. Johnson and R.M. German, in *Advances in Powder Metallurgy, P/M '91*, MPIF, Princeton, NJ, 1991.
30. J.C. Withers, S. Guha and R.O. Loutfy, *The Development of Plasma Synthesis to Produce Pre-Alloyed, Ultrafine Intermetallic Aluminide powders for Injection Molding*, Army Materials Technology Laboratory, Watertown, MA, MTL TR 91-47 (1991).

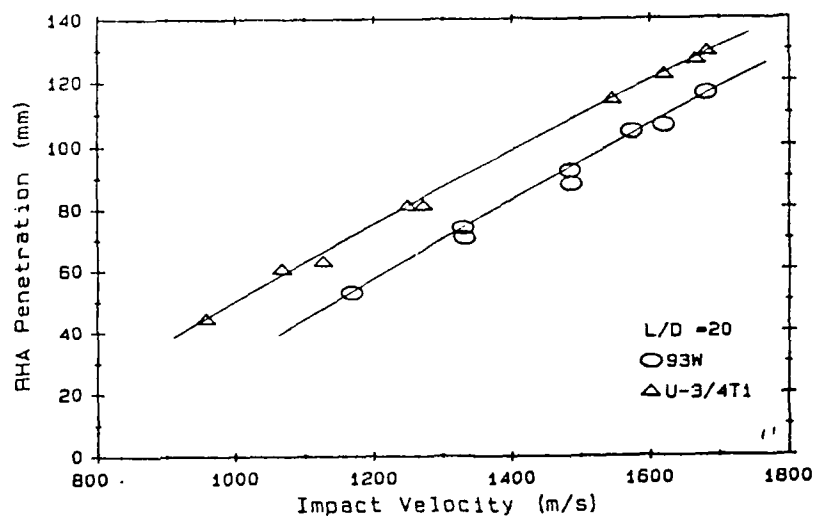


FIGURE 1 Comparison of depth of penetration of 93% tungsten alloy and depleted uranium alloy U-3/4 Ti. From Reference 1.

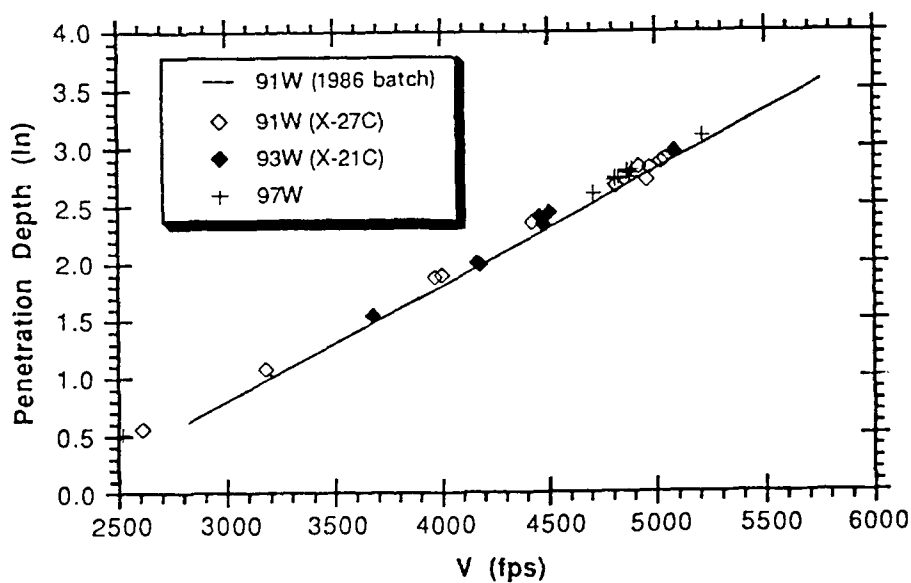


FIGURE 2 Depth of penetration of various tungsten alloys tested at AMTL. From Reference 2.

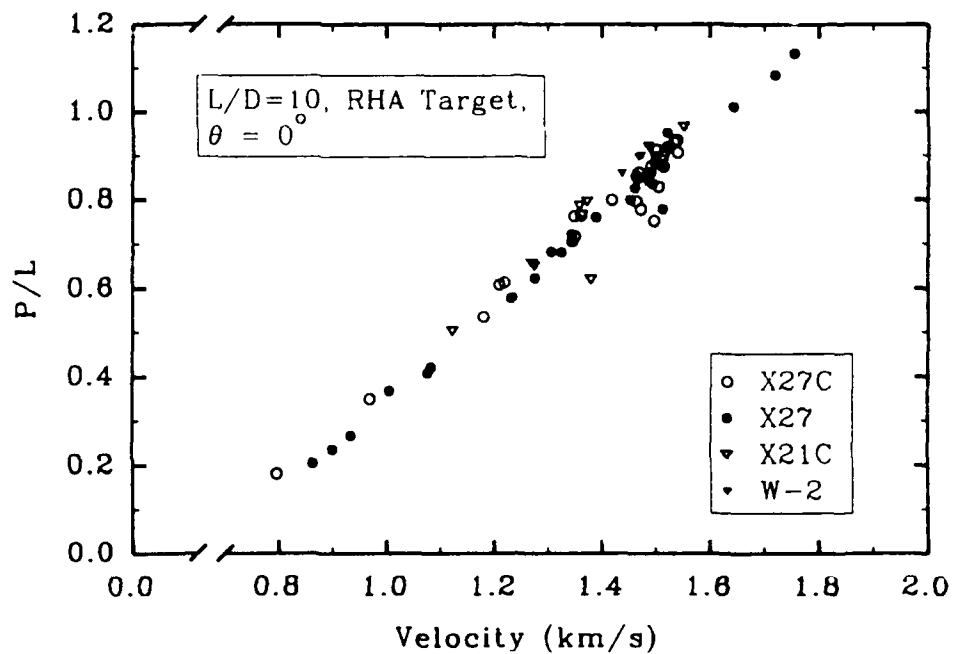


FIGURE 3 Data similar to Figure 2. From Reference 3.

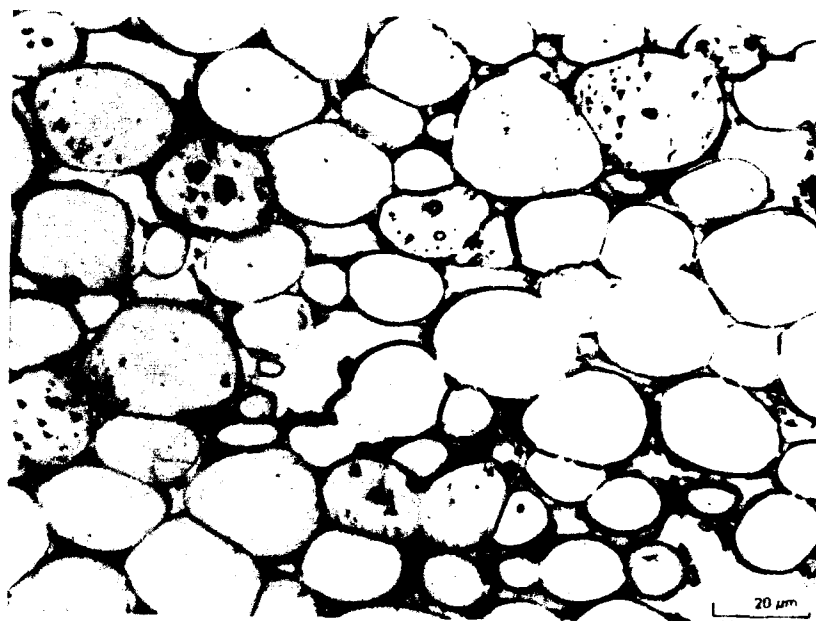


FIGURE 4 Microstructure of typical tungsten heavy alloy, 93% tungsten. Cold worked 15%.

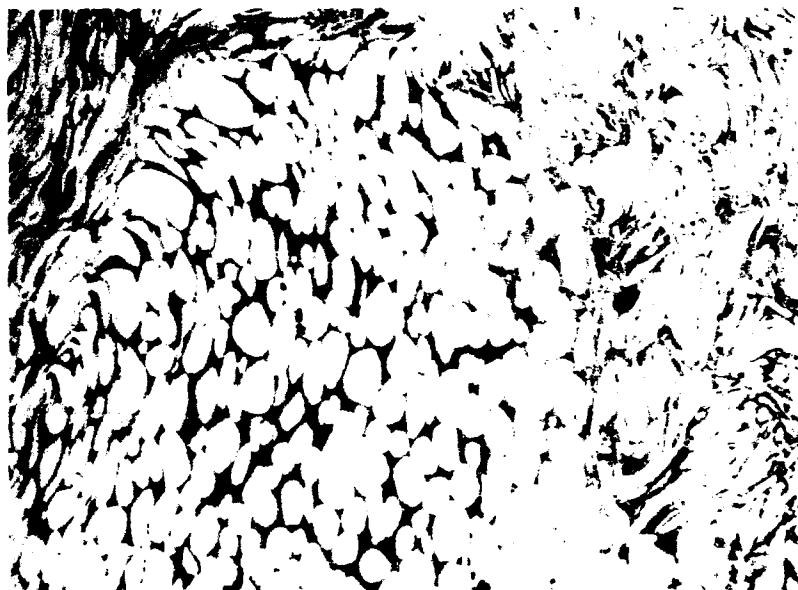


FIGURE 5 Recovered tungsten heavy alloy fragment. Figure courtesy of Messrs. E. Chin and P. Woolsey, AMTL.



FIGURE 6 Consolidated microstructure of tungsten powder CVD coated with nickel and iron. Figure courtesy of Mr. B. Williams, Ultramet.

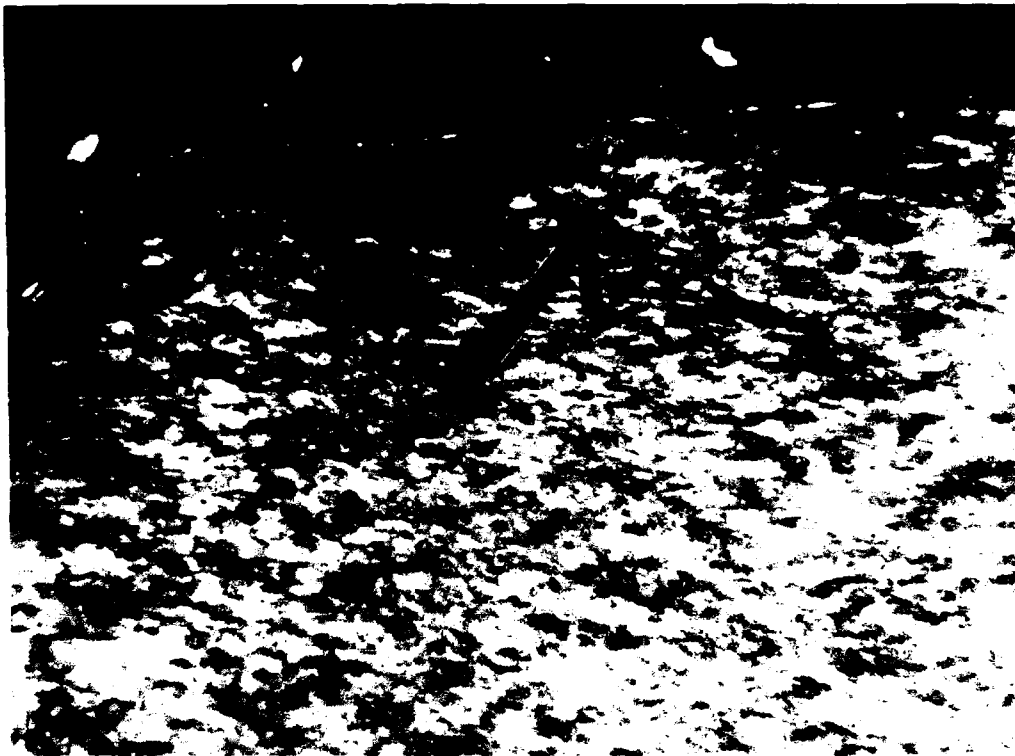


FIGURE 7 Adiabatic shear band in pure hafnium. Figure courtesy of Mr. B. Williams, Ultramet.

A Novel High Contrast Optical Modulator
in a GaAs-AlAs structure

Mitra Dutta, Dr., Hongen Shen¹, Dr., Jagadeesh Pamulapati, Dr.,
Peter G. Newman, Mr. and Wayne H. Chang, Mr.
U.S. Army Electronics Technology and Devices Lab
SLCET-ED, Fort Monmouth, NJ 07703

The performance advantages of an integrated optoelectronic approach stem from a massively parallel processing and interconnect function, wideband information processing and storage, increased switching speed and immunity to electromagnetic radiation. To meet the challenge of future systems it is important to develop integrated optics and optoelectronic devices so as to take full advantage of the photonics technology. This technology including sources, sensors, modulators, switches, optical waveguides and fibers will have a strong impact on future defense systems. The greatest impact will be felt in the fields of communications, computing and sensors. Enhanced performance of any of the elements of sources, switches, modulators, interconnects or sensors will increase the utility of optical techniques for signal processing applications.

Heterostructures and multiple quantum wells in III-V semiconductors promise fresh approach and new options for high performance electronic and optical devices required for future high speed transfer, processing and control of information. Gallium Arsenide (GaAs) based multiple quantum wells (MQW) and heterostructures have opened up new areas of science and engineering and have excited a large amount of research. With molecular beam epitaxy (MBE) and metallorganic chemical beam epitaxy (MOCVD) techniques it is now possible to prepare materials techniques it is now possible to prepare materials with a synthetic periodicity superimposed on the characteristic crystal lattice structure. The superlattice and multiple quantum well (MQW) materials have unique and optical and electrical properties that are intrinsically different from the bulk compounds. Bandgap, lattice constants and refractive constants may be tailored in III-V heterolayers to confine carriers and radiation to impose built-in fields without the penalty of surface defects and to enhance mobilities, hot

¹ GEO-Centers, Inc., Lake Hopatcong, NJ 07849. Work performed at U.S. Army ETDL.

of devices not possible before in bulk semiconductors, have been developed making use of the new physics associated with these materials and device structures. Low threshold current quantum well and wire lasers, tailored bandgap detectors, modulators and switches are being researched and developed. Most modulators are based on the quantum confined Stark effect (QCSE), the most well known of these being the logic device based on this, the self electrooptic device (SEED).

The (QCSE) is a phenomenon which arises when an electric field is applied across the plane of heterostructure superlattices. In a quantum well at zero electric field the electron and hole energy levels are defined by the well width and the electrons and holes are strongly confined in the well material layer. However when an electric field is applied the electron and holes are moved apart and their energies are altered. This has the effect of shifting the absorption resonance to lower energy as well as modulating the strength of the absorption. This shift and decrease in strength of the absorption then provides for the optical modulation of the transmission as a function of applied electric field of any light appropriate wavelength that is incident on the heterostructure. The modulator concept that we will present is based on a different mechanism, the coupling between the Γ and X valleys in a GaAs-AlAs MQW that promises to provide better contrast and operate at lower voltages than the QCSE devices.

Type II quantum well structures have been studied¹ extensively in the past several years for their interesting physics and possible device applications. In these structures, where GaAs-AlAs short period superlattices (SPS) are archetypical, the electrons and holes are confined in adjacent layers of the materials, since the direct gap GaAs layers are narrow enough so that the confined electron level in the GaAs is higher in energy than the X-valley in the indirect gap AlAs barrier. However, even when the well layers are thick, with an indirect gap barrier, effects²⁻⁵ due to the coupling between the Γ and X-levels manifest themselves. When the

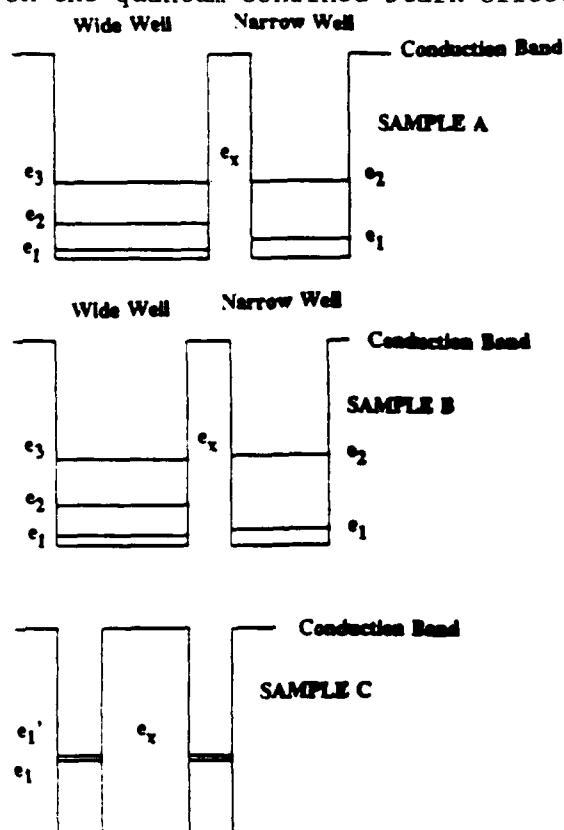


Figure 1 Energy band structures for GaAs-AlAs quantum wells.

allowed electron energy level is close to the X-valley in the barrier, the electron appears to be no longer confined, but delocalized. This has been seen previously from our measurements of exciton binding energy² which was found to be anomalously low when the electron confined level is close to the X-valley, indicating less confinement. Further evidence for this was obtained in photoluminescence excitation experiments³ where charge transfer was observed between asymmetric wells coupled via the X-valley. In some samples, the X-valley was seen to be the relevant barrier for transport⁴. In asymmetric wells coupled via the X-valley, the application of electric field⁵ showed a clear difference in the subsequent photoluminescence signal between samples that were coupled via the X-valley versus those that were not. In this paper we report experimental results on samples where the reconfinement of the delocalized electron by the application of electric field is seen more dramatically.

The samples were grown by molecular beam epitaxy in a Varian Gen II machine. The electric field experiments on the asymmetric coupled wells were done on GaAs-AlAs multiple quantum wells with eighteen narrow wells and one wide well. The wide well is used to monitor the charge transfer and the coupling between the wells via the X-valley. Sample structures with electron levels are shown in Fig. 1. Sample A has GaAs layers such that the third confined level in the wide well and the second confined level in the narrow well are both degenerate with the X-valley. In sample B the wide and the narrow wells are both made narrower so that the confined levels are pushed up and the X-valley is no longer degenerate with either electron level. In sample C, where the effect of reconfinement by electric field is seen more dramatically, the well width is such that the first electron level e_1 is nearly degenerate with the X-valley, and the e_1 level thus nearly delocalized. Schottky barriers of 80 Å of Ni-Au or 50 Å of Ti provided a transparent contact while an ohmic contact was made with the n^+ substrate. Preliminary experiments also have been performed with this structure C embedded in the intrinsic region of a surface - intrinsic - n^+ (s-i- n^+) structure⁶ where the large built in field can be modified by the intensity

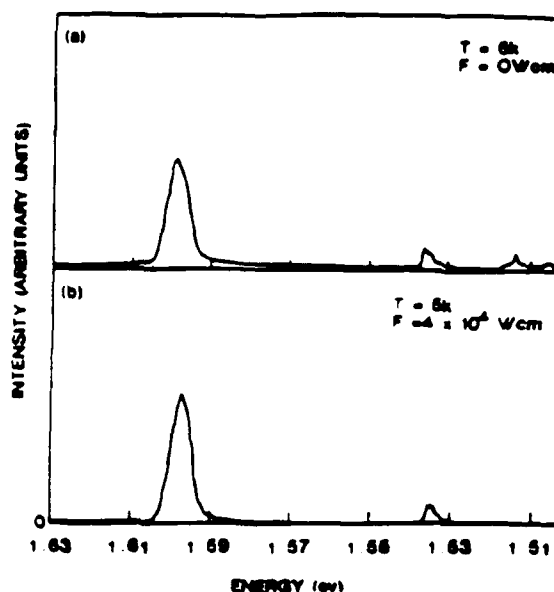


Figure 2 PL spectra for sample B; e_3 of wide well and e_2 of narrow well are not coupled with e_x ; (a) $F=0$ V/cm (b) $F=4 \times 10^4$ V/cm applied field.

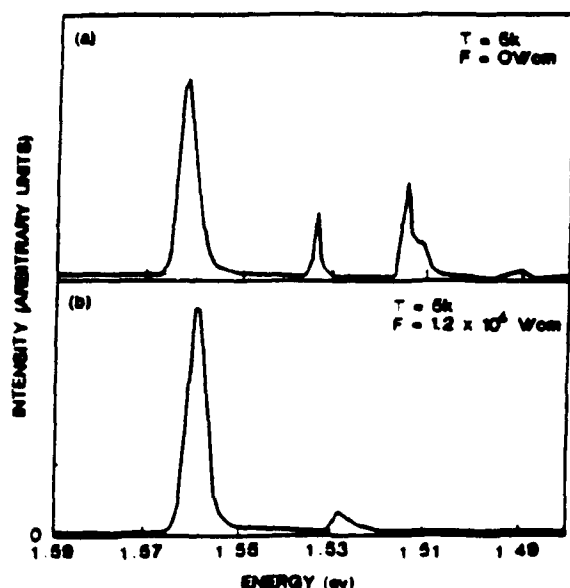


Figure 3 PL spectra for sample A; e_3 of wide well and e_2 of narrow well are coupled with e_x ; (a) $F=0$ V/cm, (b) $F=1.2 \times 10^4$ V/cm applied field.

of the light on the sample, and an externally applied electric field is no longer necessary.

All of our photoluminescence experiments are performed with the 5145Å line of an Ar^+ laser at 5K. The laser power is about 2 mW. The results of the application of the electric field on the asymmetric wells are summarized as follows: Shown in Figs. 2 and 3 are the PL spectra from samples B and A respectively. Three features are seen from each of the samples, corresponding to signal from the narrow wells, the wide well, and the bulk GaAs material. In sample B (Fig. 2), where the wells are not coupled, application of electric field ionizes the bulk GaAs exciton, and this feature disappears. The other two features from the wide and narrow wells show a red shift due to the quantum-confined Stark effect. It is important to notice that the ratio of the intensities of these two features remains the same showing that there is no coupling between the wells. This can be contrasted with sample A (Fig. 3) where there is a significant difference in the ratio of the photoluminescence intensities of the features with the application of a field; the intensity of the signal from the wide well drops more rapidly than that from the narrow wells. This is consistent with the interpretation that some of the intensity of the wide-well feature is due to electrons excited in the narrow wells which transfer into the wide well via the coupling of the X-valley in the AlAs barriers. When the coupling is removed because of the application of electric field, electrons no longer transfer into the wide well and the intensity of this feature is reduced. This effect of the confinement due to the application of an electric field is seen more convincingly in sample C. The PL spectra from sample C is shown in Fig. 4. Due to monolayer fluctuation, there are actually two such levels e_1 and e_1' , where e_1' is one monolayer less and hence of higher energy and closer to the X-valley than e_1 . The zero field spectrum shows three features of which the low energy feature at 1.702 eV is the strongest. We attribute this to the e_1 -hh₁ transition where the e_1 level is confined more strongly and is at a lower energy than the X-valley, and where hh₁ is the first heavy hole level in the GaAs well. The next feature is at an energy

12 meV higher and is the e_1' - hh_1 transitions where the e_1' electron state is nearly degenerate with the X-valley. Finally, the very weak feature at 1.725 eV is the indirect X^1 - hh_1 type II transition where the electron is the X-electron in the AlAs. This transition is made possible due to several mechanisms: (1) an overlap of the wavefunctions at the heterointerface, (2) a breakdown of the k-vector conservation by superperiodicity and disorder at the interface, and (3) possible phonon assistance. Monolayer fluctuation between a 40 Å GaAs well and a 37.5 Å well will give rise to features separated by 17 meV. However, the coupling between the Γ and X states will produce an anticrossing effect and will tend to depress in energy the e_1'

level, thus bringing it closer to the e_1 level and to the experimentally observed 12 meV between them. The electric field causes a localization of the delocalized electron and hence there is an increase in the oscillator strength as the wavefunction of the e_1' electron is confined more strongly in the well. This is analogous to Wannier-Stark ladders⁷⁻¹⁰ that have been observed when an electric field is applied to superlattice minibands. As the electric field is applied, the reduction in coupling between the Γ and X levels removes the anticrossing effect, resulting in a blue shift in the energy of the e_1' - hh_1 transition. From this blue shift, we have deduced the mixing potential between the Γ - X level to be about 3 meV, which is in agreement with the results in references 11 and 12.

It should be pointed out that the non-vertical transitions due to electrons and holes in the adjacent wells and next adjacent wells cannot be observed because of the large periodicity in our structure and the heavy effective mass of the AlAs electron. Others have observed comparable behavior in Type II short period superlattices¹⁰, but not in Type I GaAs-AlAs structures.

We have obtained similar results with the structure of sample C

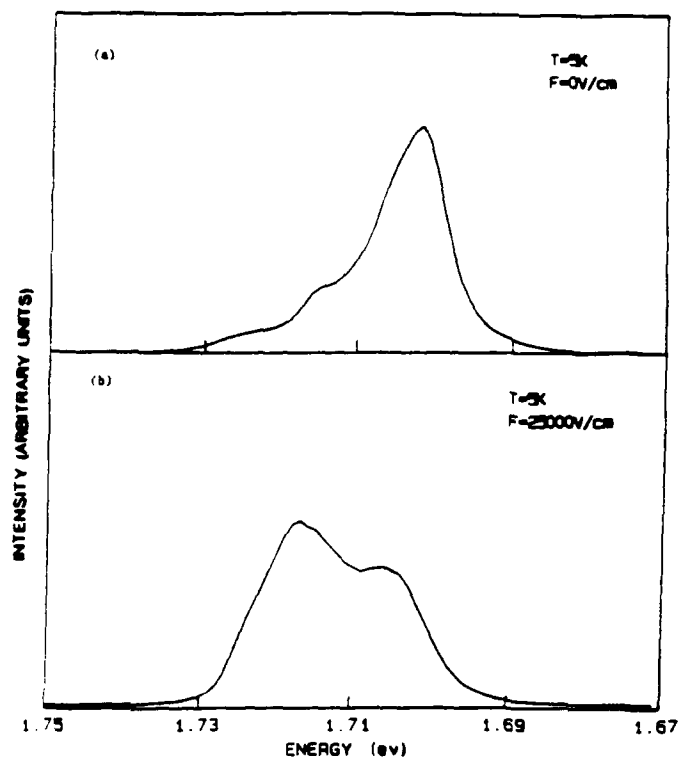


Figure 4 PL spectra for sample C; e_1' is coupled with e_x ; (a) $F=0$ V/cm, (b) $F=2.5 \times 10^4$ V/cm applied field.

embedded in the intrinsic region of the s-i-n⁺ structure, where no external field need be applied and a variation of the built-in electric field is performed by the change in intensity of the incident light. Strong dependence of the intensity of the phonon sidebands is seen as a function of electric field and more experiments with different structures are under way to understand this phenomena.

From the electric field results it can be seen how dramatically the change is in both the shift and strength of the transition at a lower electric field. These initial results compare very favourably with the QCSE based structures.

We present experimental data on GaAs-AlAs multiple quantum wells where the confined electron level is initially delocalized due to mixing between the Γ and X electron levels. The application of an electric field reduces this coupling and reconfines the electron in the GaAs layer. Two phenomena are associated with this: (1) an increase in the oscillator strength, and (2) a blue shift due to the reduction of the anticrossing between the Γ and X states; the reduction in the anticrossing is consistent with the value of the mixing potential as measured by other groups. This phenomenon can be utilized to produce a higher contrast, lower operating voltage optical modulator.

1. B.A. Wilson, IEEE J. Quant. Electron QE-24, (1988) 1763 and References cited therein.
2. D.D. Smith, M. Dutta, X.C. Liu, A.E. Terzis, A. Petrou, M.W. Cole and P.G. Newman, Phys. Rev B 40, 1407 (1989).
3. M. Dutta, D.D. Smith, P.G. Newman, X.C. Liu and A. Petrou, Phys. Rev B 42, 1474 (1990).
4. M. Dutta, K.K. Choi, and P.G. Newman, Appl. Phys. Lett 55, 2429 (1989).
5. M. Dutta, H. Shen, D.D. Smith, K.K. Choi, and P.G. Newman, to be published in Surface Science (1992).
6. H. Shen, M. Dutta, L. Fotiadis, P.G. Newman, R.P. Moerkirk, W.H. Chang, and R.N. Sacks, Appl. Phys. Lett. 57, 2118 (1990).
7. E.E. Mendez, F. Agullo-Rueda, and T.M. Hong, Phys. Rev. Lett. 53, 2632 (1988).
8. J. Bleuse, P. Voisin, M. Alloron, and M. Quillce, Appl. Phys. Lett. 53, 2632 (1988).

9. P. Voisin, J. Bleuse, C. Bouche, S. Gaillard, A. Alibert, and A. Regreny, Phys. Rev. Lett. 61, 1639 (1988).
10. A.J. Shields, P.C. Klipstein, M.S. Skolnick, G.W. Smith, and C.R. Whitehouse, Phys. Rev. B42, 5879 (1990).
11. M.H. Meynadier, R.E. Nahory, J.M. Worlock, M.C. Tamargo, J.L. de Miguel and M.D. Sturge, Phys. Rev. Lett. 60, 1338 (1988).
12. N.J. Pulsford, R.J. Nicholas, P. Dawson, K.J. Moore, G. Duggan and C.T.B. Foxon, Phys. Rev. Lett. 63, 2284 (1989).

Imagery Exploitation System/Balanced Technology Initiative

Frederick H. Esch
U.S. Army Topographic Engineering Center (TEC)
Fort Belvoir, Virginia 22060-5546

1. INTRODUCTION

The Imagery Exploitation System/Balanced Technology Initiative (IES/BTI) is a automated first phase near-real-time image exploitation system to support Army Corps Intelligence Electronic Warfare situation development, target development and target acquisition. IES/BTI's speed, efficiency, ability to perform force analysis, and ability to use low-resolution imagery will allow timely exploitation of imagery covering hundreds of square kilometers allowing commanders to see much larger portions of the battlefield.

IES/BTI exploits synthetic aperture radar (SAR) and is integrating infrared exploitation capabilities. IES/BTI locates the presence and type of military units (artillery, armor, etc.) within imagery and prepares either graphic overlays of results and/or an ASCII file of results. Currently, experiments using IES/BTI results as input to Flying Carpet (Simnet) and using IES/BTI results as input to the Demons imagery displays at TEC and the imagery analyst (IA) workstation at the imagery analyst school at Ft Hauchuca are being conducted. IES/BTI is also being considered for incorporation in the Imagery Processing and Dissemination System (IPDS) architecture at Corps.

The focus of this paper is the Artificial Intelligence (AI) aspects of IES/BTI, the conditions that influenced the chosen implementation and the lessons learned in evaluating IES/BTI in two military theaters.

ESCH

Section 2 provides a background on IES/BTI functions, and the conditions, constraints, and requirements placed on IES/BTI. Section 3 briefly addresses the functionality included in IES/BTI. Section 4 describes the lessons learned in the process of building, testing and migrating an AI system to a second military theater.

2. BACKGROUND

This section provides a brief background on the role IES/BTI is designed to play in the image analysis process and the conditions, constraints and requirements placed on IES/BTI.

The envisioned role of IES/BTI in the Image Exploitation Process is to automatically perform exploitation before an IA receives the imagery (see Figure 1). IES/BTI performs vehicle detection, force aggregation, force typing, false hypotheses rejection and merging of various evidence sources. The IA is presented the IES/BTI results along with the imagery. Guided by the results graphic (see Figure 2), the IA can quickly focus on the potential forces and can efficiently eliminate major areas of search. With this tool much larger amounts of imagery can be processed by the Corps level analyst and the analyst's task shifts away from vehicle level detection to unit level analysis/verification.

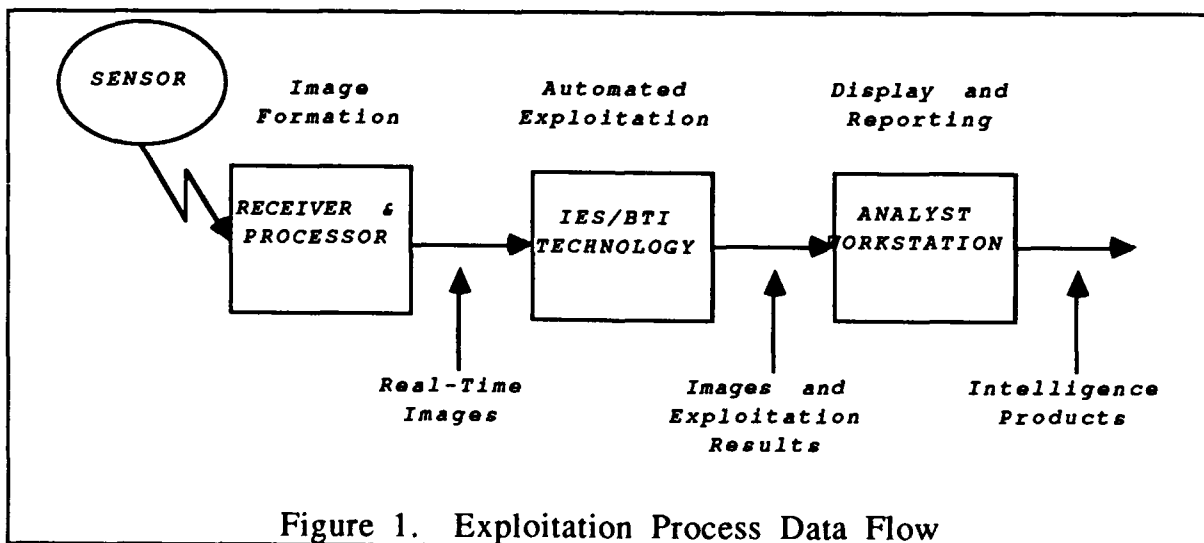


Figure 1. Exploitation Process Data Flow

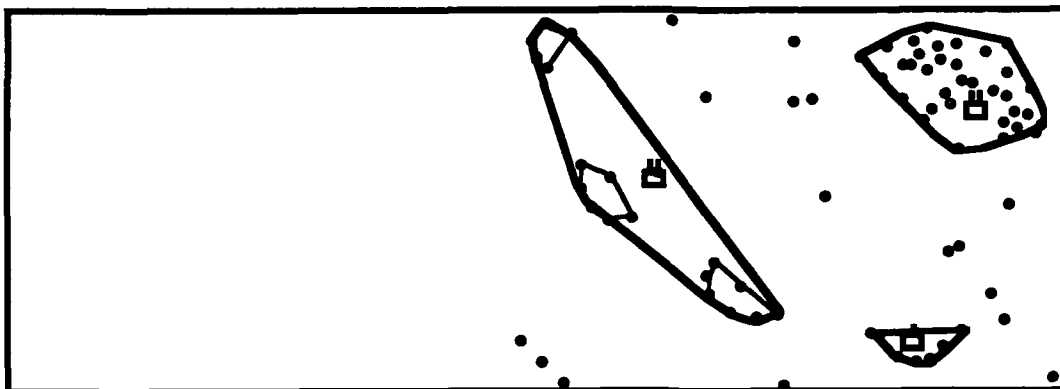


Figure 2. Example results graphic. Potential vehicles are shown as dots. Companies and battalions are shown as hulls and icons.

To illustrate how an IA uses IES/BTI results an example image exploitation event is presented. Gil, an IA, has been asked to exploit an image. Gil calls up the image and the IES/BTI results graphic on his workstation. Gil sees the result graphic, Figure 2, draped over the image. Gil notices the left side of the image contains no activity and being a pessimist quickly scans the left side of the image. Gil knows that the results graphic contains over 90 percent of the vehicles present so unless something catches his eye, he can feel comfortable ignoring the left side of the image. Satisfied the left side is empty, Gil focuses his attention on the right side of the image. Gil then verifies the forces presented on the results graphic. Gil asks himself are all the units real? Did the machine miss any units? Is the lone company part of either of the battalions? Believing that the lone company is part of one of the battalions, Gil adjusts the interpretation to Figure 3 and generates his report.

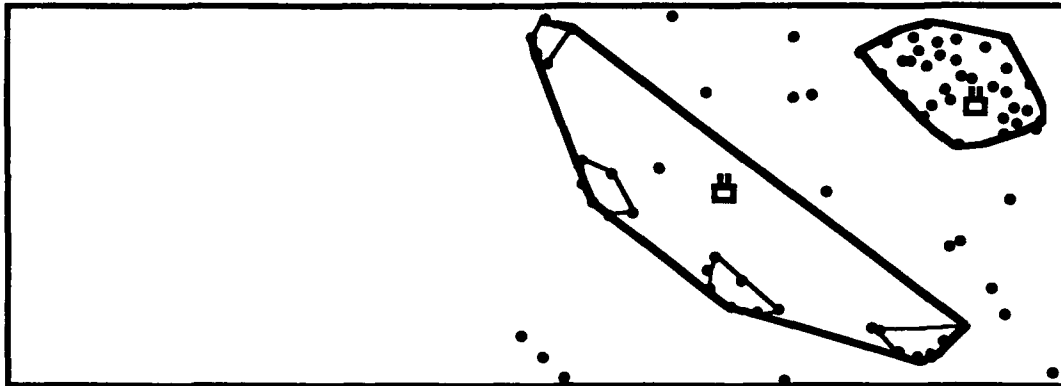


Figure 3. Gil's results.

IES/BTI's system requirements are oriented for a tactical environment. IES/BTI will be used by IAs rather than computer scientists so the results and supplemental information must be clear. IES/BTI will be used in many military theaters and will use standard DMA products as a basis for terrain reasoning. IES/BTI requirements include timing, accuracy, multi-sensor considerations, software standards, compliance with commercial off the shelf policy (COTS) and theater of operation portability requirements. Additionally, IES/BTI will be constrained to two 19 inch racks of computer equipment. Top level requirements are given in Table 1.

Table 1. Top Level IES/BTI System Requirements

Requirement
Provide multi-sensor capability (SAR, IR)
Complete processing of image in less than five minutes for images up to 100 sq nm
Process in a operational (tactical) setting
Provide 100 m vehicle location accuracy
Provide graceful migration between theaters of operation
Use DMA products for terrain information

The system requirements, especially coverage, speed, theater portability, and data sources, put constraints on the AI methods and

ESCH

representations that IES/BTI could use. Being designed as a field deployable system also constrains AI methodologies by not confining the problem and creating indirect requirements of being well behaved and having results that degrade gracefully. Thus it is possible for IES/BTI to be asked to analyze imagery from a new military theater, against a new force in a matter of days! Coping with this scenario presents a difficult task for an AI based system.

3. IES/BTI FUNCTIONALITY

IES/BTI performs five functions: vehicle detection, terrain reasoning, force aggregation, force evaluation/conflict resolution and force typing. Functions within IES/BTI have three notable characteristics: 1) functions are actually large sets of evidence-gathering actions that have been abstracted to a function; 2) functions may be split between several software components that may be widely separated in the reasoning chain; 3) boundaries between functions blur at higher levels of reasoning especially when using multiple sources of evidence. These characteristics are important as they allow us to meet several of our requirements and influence the AI methodology chosen. A brief discussion of each function and Bayesian inference is given below.

An inference system is needed to knit the functions' results together and arrive at an answer. Within IES/BTI we chose Bayesian inference using a Bayes network of hypotheses and Pearl's algorithm for belief propagation.

Bayesian inference has several properties which we found beneficial: a basis in probability, a well defined method of propagating evidence, a single belief value for a hypotheses, a hierarchical network that relates well to force structure and allows efficient management of hypotheses, a model that estimates the utility of future actions, and independence of belief and order of evidence aggregation.

Within IES/BTI, we use the hierarchy of the Bayes network to correlate with force echelon. Each Bayes node contains the current set of conflicting hypotheses (at the same echelon) with each hypothesis

ESCH

linked to its supporting evidence, child nodes and parent nodes. The number of hypotheses within a node is controlled by pruning hypotheses with beliefs below a minimum belief threshold and by pruning the weakest hypotheses when the number of hypotheses within the node exceeds a threshold¹. Controlling the number of hypotheses is performed to limit combinatoric explosion and allow quicker processing.

Vehicle detection within IES/BTI is based on window operators and does not have a AI implementation. However, it should be understood the role of vehicle detection is to detect greater than 90 percent (>90%) of the vehicles present regardless of false alarm rate. In the speckled SAR environment more than 50 percent (>50%) of detections may be false alarms! We prefer a high detection rate rather than a low false alarm rate as we do not want the IA to have to search the image for missing vehicles and therefore defeat the purpose of giving him an IES/BTI! Secondly, when detection and false alarm rates are lowered, entire units may be lost or rejected in early processing causing the analyst to search the imagery for missing units.

Therefore, other functions must perform rejection of false alarm detections, perform rejection of false unit hypotheses and reason about units that may include false vehicle detections. This means most reasoning in IES/BTI is geared toward separating the real units (with noise) from the phantoms.

Terrain reasoning is performed at several stages within IES/BTI. The coarser reasoning of rejecting detections in prohibited (No-Go) areas and mobility/emplacement value calculation is based on military order of battle, terrain analysis and heuristic models. The coarse generic reasoning is precompiled and saved as encoded data layers for each military theater. At run-time local operations calculate a generic (non-unit specific) terrain likelihood ratio value with the emphasis on providing a rank ordering of clusters based on the mobility/emplacement value of the terrain underlying the cluster. The

¹Null (no unit) hypotheses are never culled.

ESCH

importance of rank ordering clusters is to allow subsequent processing to focus on the most promising clusters and to start rejecting clusters with low belief values.

Later stages of terrain reasoning are divided into two categories: reevaluation of mobility/emplacement value or as expert system rules utilizing conceptual terrain objects (hill, city, avenue of approach, etc). Reevaluation of mobility/emplacement potential can be performed to obtain mobility/emplacement potential for a specific type of unit (such as artillery) or as part of an inter-force/terrain evaluation. Force/mission/terrain evaluation and inter-force/terrain evaluations are often encoded as expert system rules. Within IES/BTI terrain expert system rules are generally Army and military theater specific and terrain objects used by the expert system are regionally defined.

Force aggregation consists of two types of actions: an initial clustering phase where vehicle detections are clustered into potential company or battalion sized units and a second phase aggregating (hypothesized) units into a higher echelon unit. Clustering attempts to form all real units while limiting the number of non-real units. Clustering eliminates built-up areas and high speckle areas by density screening detections, and then clustering aggregates detections that are related both spatially and by detection belief. Clusters are evaluated on image derived evidence and the underlying terrain (mobility/emplacement value). Clusters form the basis of the initial hypotheses formed in the Bayes network.

A natural method of force aggregation is inherent in Bayesian inference using a Bayes network. As part of building a Bayes network, Bayes nodes are created and linked hierarchically. Within IES/BTI, we use the hierarchy of the Bayes network to correlate with force echelon. Force aggregation is simply building the next level of the network and applying the appropriate force evaluation actions.

Force evaluations are actions that can generate new hypotheses, add evidence to hypotheses and produce or alter Bayes network evidence nodes. Force evaluations evaluate hypotheses features and/or beliefs

ESCH

with hierarchical models of forces. The evaluated features often include features of the child hypotheses. Double counting of evidence is avoided by adding or altering evidence nodes at the appropriate echelon/Bayes network level and propagating all other evidence up or down the Bayes network. Evaluation criteria are stored in a knowledge base organized hierarchically by force level, force type and evidence granularity. Evidence gathering actions evaluate the closeness of data supporting the hypotheses to the force models using statistical metrics and/or expert system rules. Some rules are complex dynamically adapting a static force model to the local terrain and threat situation. Expert system and other high cost rules are applied only if more coarse evidence evaluations and action utility calculation support proceeding with the action.

System level conflict resolution consists of deciding which actions to perform and deciding when to stop processing and report results. System level conflict resolution is supported by the inherent ability to monitor the state of the Bayes network and therefore propose actions based on that state and the ability to calculate the utility of the proposed actions.

Within the Bayes network, hypotheses conflict resolution is not necessary except to ensure efficiency. All conflicting (mutually exclusive) hypotheses can be kept in a single Bayes node and all hypotheses can be carried indefinitely. This feature combined with the ability to estimate the utility of future actions and the ability to monitor the state of the Bayes network, allows timely low-cost belief-based dynamic hypotheses pruning.

Force typing actions are actions which attempt to identify a force's type (armor, artillery, etc) and thus refine the hypotheses. Figure 4 shows the hierarchical organization of force types used within IES/BTI. Competing hypotheses differing in force type exist within the same Bayes node. Force typing actions are quite variable and try to seize on the distinction between the force types. Force typing actions often generate new competing hypotheses within the Bayes nodes and may

internally generate and prune hypotheses as part of the action. Force typing are generally hierarchical actions.

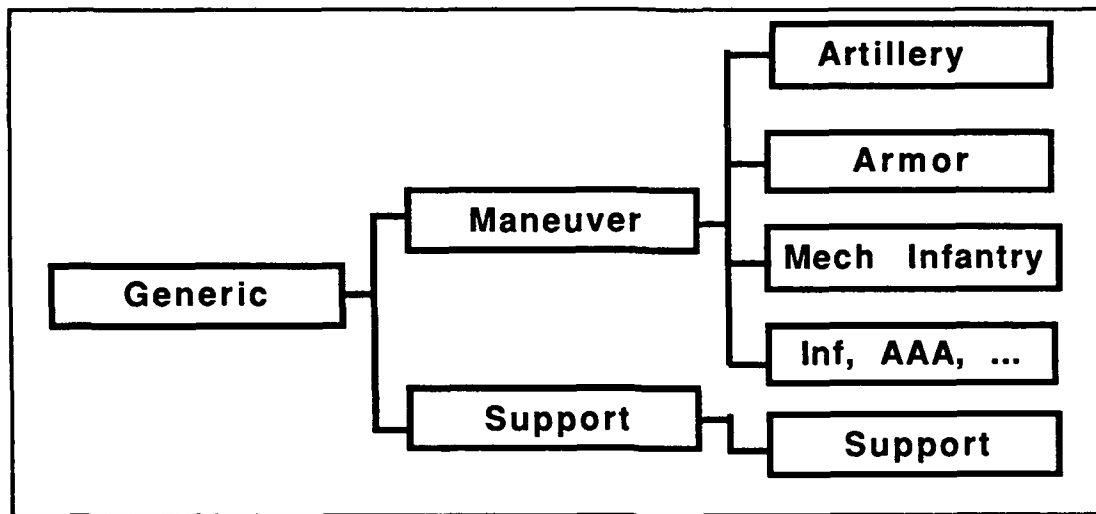


Figure 4. Hierarchical organization of unit type within IES/BTI.

Methods for identifying unit type within IES/BTI are: using vehicle level information, using deployment and formation information, using terrain considerations and using intelligence cues.

4. LESSONS LEARNED

Within the IES/BTI development team there have been several ideas that influenced the success of IES/BTI that are applicable to a wide range of AI applications. These ideas proved their value in two military theaters, Europe and South West Asia, even though the scenarios are much different. These winning ideas are given in Section 4.1.

There were also ideas within IES/BTI that fell short of our expectations. These ideas failed to mature in our environment and we provide observations or suggest alternate methods. Some of these ideas may work out for you and time and technology will eventually bring

about their success. Nevertheless, our failure may be your savings so they are included. Our ideas that fell-short are given in Section 4.2 and include: image-to-map-registration, formation pattern matching and automatic target recognition.

4.1 WINNING IDEAS

Within the IES/BTI development team there have been several ideas that worked out well and are applicable to a wide range of AI applications. The winning ideas are flexibility, reasoning in real-world coordinates, Bayesian inference and hierarchical reasoning.

Flexibility is the single most valuable asset for a military AI system. A comparison of the relative evidence value provided by IES/BTI functions for the European and the South West Asia Theaters demonstrates this point, see Table 2.

Table 2. Relative Value of Evidence Sources by Military Theater

Function	Europe	S W Asia
Detection	Low	High
Terrain Reasoning	High	Low
Force Aggregation		
Clustering	Medium	High
Up Echelon	High	Medium
Force Evaluation	High	Medium
Force Typing	Medium	Low

As shown in Table 2, the relative value of evidence sources are not constant from theater to theater. A system that requires consistent behavior may perform poorly outside the theater used in designing it. Secondly, flexibility is required as expected behavior and reality are not exact matches and dynamic adapting to doctrine must occur.

Abstracting to high-level functions increases flexibility in definition and implementation which is crucial when changing military theaters,

ESCH

adding sensors or adding additional evidence sources. Using knowledge bases also promotes theater portability as does parameterizing any heuristics. Overall the less functions are hard-coded the more portable they become.

Symbolic reasoning in real-world coordinates (latitude/longitude or Universal Transverse Mercator (UTM)) allows a AI application to be sensor independent. Leaving the image domain as soon as possible allows other processing to be sensor independent and maximizes symbolic reasoning and code reuse. Within IES/BTI, adding sensors only requires adding the appropriate detection and registration code.

Neural networks, Dempster-Shafer, cluster analysis and decision trees were among the methods considered for use in IES/BTI. However, no other method contained the combined features, power and flexibility of Bayesian inference. Bayesian inference insensitivity to small errors in belief and the ability to re-calibrate evidence belief values are needed in real-world applications. An AI system without the ability to economically recalibrate evidence may be left fighting the last war. Furthermore, Bayesian inference behaves well regardless of activity level. Bayesian inference neither forces you to keep hypotheses nor forces you to prune hypotheses, allowing you to correctly reflect the activity level as shown in Figure 5. Other useful features such as: a probabilistic basis, a well defined method of propagating evidence, a single belief value for a hypotheses, hierarchical organization, utility estimation of actions and independence of belief, and order of evidence aggregation were pointed out in Section 3. Not all applications will benefit from Bayesian inference, but systems that require symbolic logic across several evidence domains, have potential combinatorial problems, need to be portable across military theaters or that can take advantage of hierarchical reasoning may benefit from Bayesian inference.

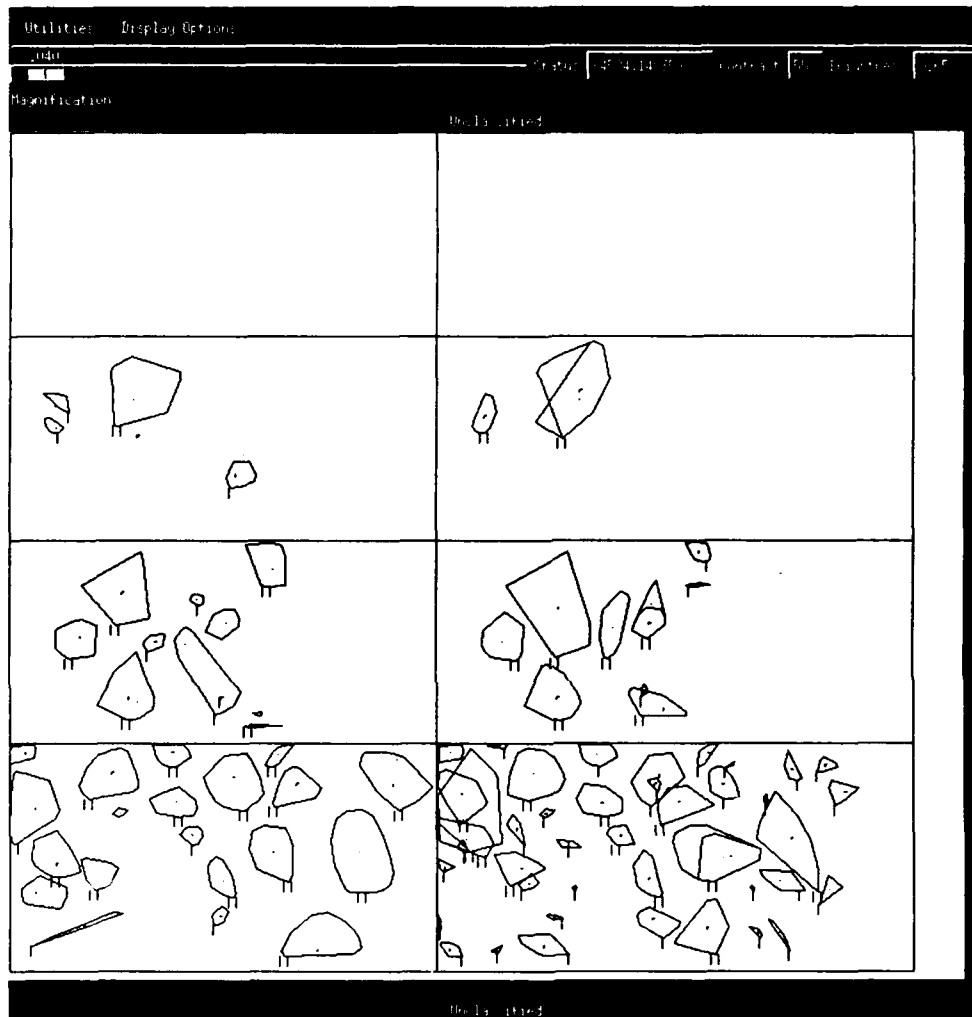


Figure 5. Truth (left) compared with IES/BTI results (right) for four different levels of unit activity (strength). Hulls with two "legs" are the convex perimeter of a battalion. Hulls with one "leg" are the convex perimeter of a company. Hulls with no "legs" are the convex perimeter of a element. Note for IES/BTI results overlapping hulls indicate competing hypotheses with similar belief and the IA is left to resolve the unit extent.

Hierarchical reasoning was another big winner. Hierarchical reasoning matches well with Bayesian inference and is necessary in near-real time systems to quickly produce results as exhaustive evaluation is too costly. Hierarchical reasoning is needed in the high speckle SAR environment where applying detailed models against hypotheses without prior coarse evaluations leads to phantom matches. Hierarchical reasoning avoids rejection of true units by allowing the hypotheses to be kept and confirmed at a coarse level even if it fails to successfully refine into a more detailed hypotheses, while allowing pruning of poor hypotheses as evidence against them builds up.

4.2 IDEAS THAT FELL SHORT IN IES/BTI

Our ideas that fell-short include: image-to-map registration, formation pattern matching and automatic target recognition.

Image-to-map registration failed as it misuses map derived data. Maps were never meant to serve as a basis of photogrammetric control but people need to register imagery to the earth, and maps are often the only terrain product available. Maps contain "cartographic licence" which alters the metric relationships between features, thus introducing registration errors. Cartographic licence is needed in map making to produce clear maps. Thankfully, DMA is producing image derived products (video point positioning data base (VPPDB) and some interim terrain data (ITD)) which do not contain the cartographic licence inherent in maps. Anyone contemplating image-to-map registration might pursue image-to-VPPDB as a more likely candidate for success.

Formation pattern matching did not meet our expectations in both military theaters for two different reasons. In Europe, forces moved a lot and didn't reflect classical formations (V, line, six up one back, etc) when in movement, halted near a road or in assembly. In South West Asia all units tended to have similar patterns or no pattern regardless of forces present. Finally, regardless of theater, missing vehicles within the pattern greatly reduces the value of the "match". However, pattern matching is effective for units such as artillery, where physics encourages them to form regular patterns.

ESCH

Automatic target recognition (ATR) also fell short of expectations. ATR promises accurate unit typing and accurate unit verification. IES/BTI does not currently use an ATR for two main reasons: ATR is currently too slow to meet IES/BTI timelines and ATR rarely discriminates against all the vehicle classes/types needed in a real-world battlefield environment. As ATR matures IES/BTI will surely include them.

5. REFERENCES

1. J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, (Springer-Verlag, New York, 1985).
2. A. Klinger, H. Thomas, and B. Haddock, *IES/BTI Concept of Operations (Draft)*, Imagery Exploitation System / Balanced Technology Initiative Internal Document, (1991).
3. T. S. Levitt, et al., *IES/BTI System Design Document*, Imagery Exploitation System / Balanced Technology Initiative Internal Document, (1991).
4. A. M. Mood, F. A. Graybill, and D.C. Boes, *Introduction to the Theory of Statistics*, McGraw Hill, New York, (1974).
5. J. Pearl, *Artificial Intelligence* 28, 9, (1986).

Observation of rf-Induced Failures in Silicon Bipolar Junctions Using Synchrotron X-Ray Topographic Imaging

C. Fazi

U.S. Army Harry Diamond Laboratories
2800 Powder Mill Road
Adelphi, MD 20783

M. Dudley, S. Wang, and D. Gordon-Smith
Department of Materials Science & Engineering
SUNY at Stony Brook, Stony Brook, NY 11794

1. Introduction

The U.S. Army is concerned about the survivability of military systems employing semiconductor devices in high transient radio frequency (rf) environments such as high-power microwaves (HPM) and nuclear-generated electromagnetic pulse (EMP). Understanding the nature and causes of the damage induced in electronic semiconductor devices by transient electromagnetic fields is an issue of central importance to assessing the long-term survivability of such devices and to hardening them. This damage is associated with drastic changes in the electrical characteristics of the devices, which occur during uncontrolled breakdown processes.

Apart from detailed electrical characterization of the devices before and after breakdown, investigations of the nature of the resulting damage have been largely limited to optical microscopy and scanning electron microscopy (SEM) studies (for example, see Ricketts et al [1]). By nature, optical and SEM studies can reveal modifications only in the surface morphology of the metallization and are not sensitive to subsurface damage. SEM of cross-sectionally bevelled specimens can reveal some information (for example, regarding distribution of chemical species), but it cannot give a reliable indication of the distribution of damage throughout the epilayer region. Transmission electron microscopy (TEM) is generally not routinely applicable in such studies for two reasons: (1) the technique is destructive by nature, since thin specimens must be prepared (potentially removing some of the damaged regions distributed throughout the thickness of the device) and (2) the technique has a

very small field of view, making it very difficult to prepare specimens specifically containing the localized damaged areas. In order to reveal such subsurface damage, a nondestructive technique capable of high strain resolution is desirable. X-ray diffraction topography provides a potential solution, although in either transmission or conventional Bragg geometries, sensitivity to distortions located only a few micrometers below the surface of a crystal is very limited. However, employment of grazing Bragg-Laue geometries enhances the near-surface sensitivity of the technique. The versatility of the technique in these geometries is greatly enhanced by the application of synchrotron white radiation, with its tunability and high brightness (see Yao et al [2]).

The use of synchrotron x-ray topography to image the damage created during rf-induced failures in semiconductors has already been proven to be an important diagnostic tool in understanding rf hardening techniques and failures in semiconductors (see Dudley et al [3]). Unlike conventional x-ray sources, the synchrotron radiation source can produce area-filling beams of highly collimated white x-rays. The large beam area, coupled with high intensity, make it possible to record many images simultaneously (comprising a Laue pattern) with exposure times of just a few seconds. Transmission geometries enable bulk strain distributions in the crystal to be revealed, while the use of grazing-incidence Bragg-Laue geometries reveals detailed information of local strain distributions as a function of depth.

This paper is a preliminary report of results of a synchrotron x-ray diffraction topographic investigation of damage deliberately induced in *pn* junctions on silicon by high-power microwave sources and video pulses of 3 μ s and 200 ns duration. Information obtained from x-ray topographs is compared to that obtained from optical micrographs. Observed differences are discussed, and the superior nature of the information afforded by the x-ray technique is demonstrated.

2. Experimental Results

Specimens used in these experiment were fabricated at the U.S. Army Harry Diamond Laboratories. Samples consisted of arrays of diodes of different shapes and sizes on 2-in. diameter (111) silicon wafers. The wafers consists of a 12- μ m-thick *N*-type epilayer on a 0.01-in. thick *N*⁺-type substrate, with an epilayer doping level of 10^{15} impurities/cm³. The use of wafers rather than individually packaged devices facilitated the testing and subsequent simultaneous imaging of statistically significant numbers of devices which had undergone identical processing. A cross-sectional view for a *pn* junction is given in figure 1. We induced breakdown in selected 300- μ m-diameter circular junctions by direct injection of a single 3- μ s-duration or 200-ns-duration unipolar pulse of reverse polarity. This was achieved

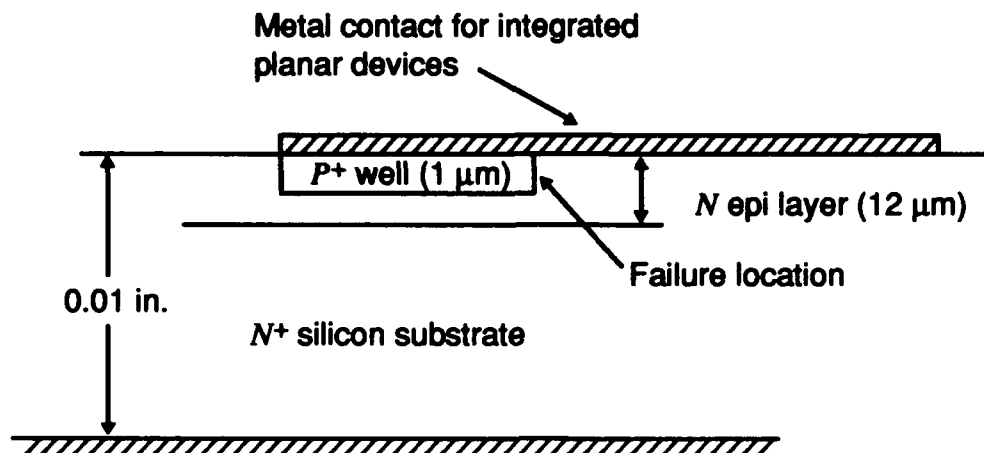


Figure 1. Cross-sectional view of a *pn* junction used in integrated planar devices. Contacting P^+ well with a metal line induces failures at this interface, as arrow indicates failure location. Discrete semiconductor devices usually do not rely on this type of contact, hence may be less susceptible to failure.

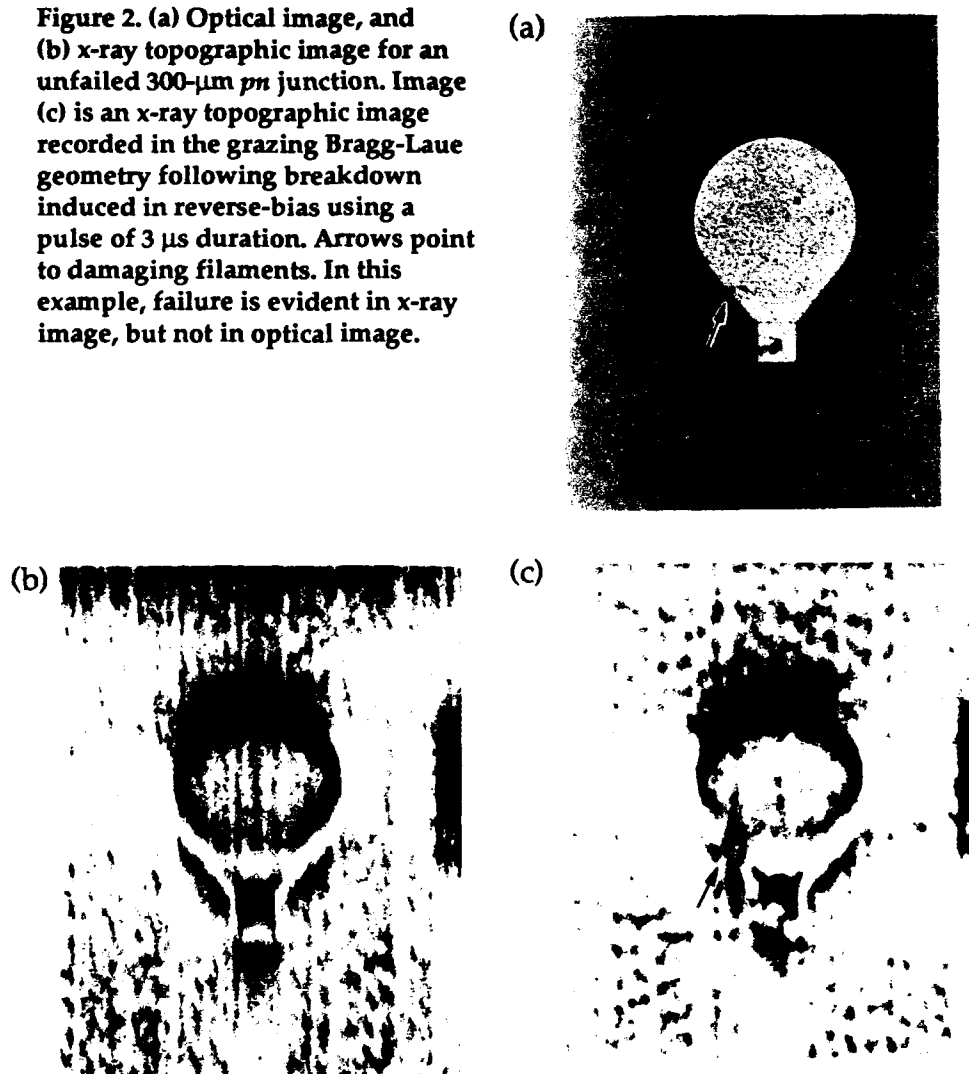
with the use of a coaxial probe technique developed by Fazi [4]. The amplitude was increased until permanent changes in the breakdown voltage were observed on a curve tracer. HPM breakdown was induced in selected triangular junctions of the same surface area as the circular ones. The purpose in switching topology for the HPM test was to induce failures with the lowest possible pulse amplitude, since high electric field at the corners would induce failures there. It did not; instead, it was surprising to observe that the induced failures for both topologies, driven by HPM and video-pulse-type waveforms (narrowband and wideband) occurred only between the metallic interconnect and the *pn* junction, as indicated in figure 1 by an arrow showing the failure location. For discrete devices, the failure location has been observed to occur randomly around the P^+ well, and the failure levels are proportional to the junction volume. If the failure is attributed only to an edge effect, the modeling is no longer a three-dimensional issue. The correct physics must be taken into account when setting the boundary conditions for the damage modeling codes. Based on this preliminary observation, one may wonder if planar technologies (including very-high-speed integrated circuits and microwave monolithic integrated circuits) may be inherently more vulnerable to rf burnout than discrete vertically structured devices.

We carried out white-beam transmission and variable-penetration-depth topographic imaging in grazing Bragg-Laue geometries using the white-beam camera on the Stony Brook Synchrotron Topography Beamline (X-19C) at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. The variable-

penetration-depth grazing Bragg-Laue imaging was carried out using methodologies that can be found in the literature [2,5-8].

Figures 2 and 3 show optical and x-ray topographic images recorded in the grazing Bragg-Laue geometry from two 300- μm -diameter circular diodes following breakdown induced by pulses of 200 ns and 3 μs duration, respectively. Note that several features are visible in the second optical picture that can be interpreted as failures. Only after comparing with the x-ray images can one determine that the feature at the interface edge is due to a damaging filament (indicated by an arrow).

Figure 2. (a) Optical image, and (b) x-ray topographic image for an unfailed 300- μm *pn* junction. Image (c) is an x-ray topographic image recorded in the grazing Bragg-Laue geometry following breakdown induced in reverse-bias using a pulse of 3 μs duration. Arrows point to damaging filaments. In this example, failure is evident in x-ray image, but not in optical image.



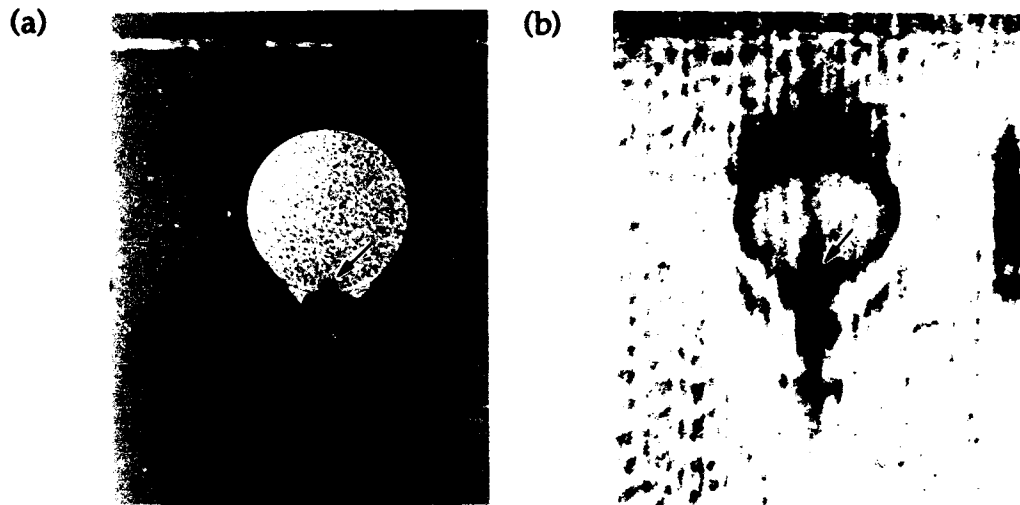


Figure 3. (a) Optical image and (b) x-ray topographic image recorded in the grazing Bragg-Laue geometry following breakdown induced in reverse bias using a pulse of 200 ns duration. Arrows point to damaging filaments. In this example, failure is evident in both types of image.

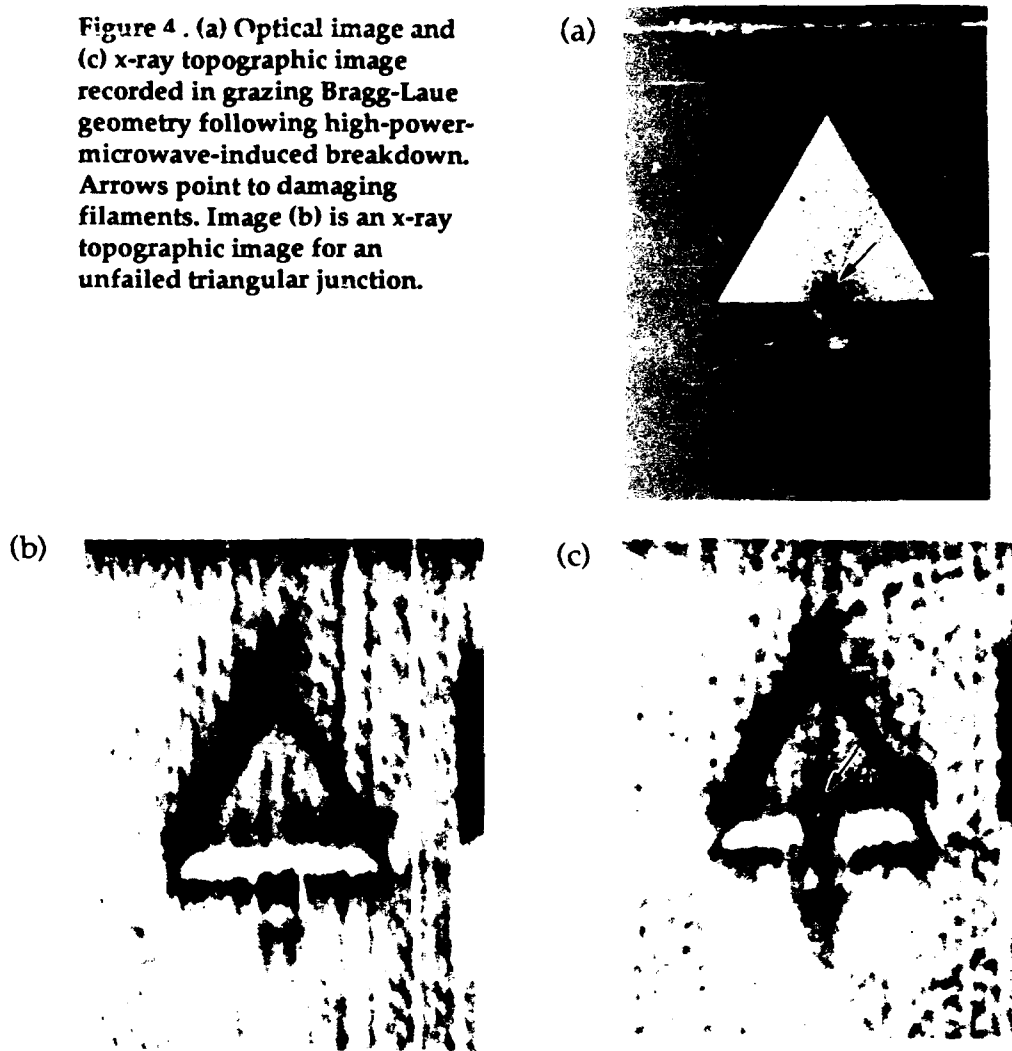
Figure 4 shows similar optical and x-ray topographic images recorded from a triangular diode of equivalent surface area following HPM-induced breakdown. The purpose of changing to a triangular topology is to attempt to induce failures at the sharp corners due to high field-induced breakdown. Again the failure occurred at the metal-to-junction interface (indicated by an arrow).

3. Conclusion

It appears, from the limited data taken so far on planar epitaxial devices, that rf transients of either narrow-band or wide-band type may induce filamentary failures at the interface where the metallic line contact and the junction meet, as shown in figures 2(a) and 3(a) by the arrows. This poses a dilemma when one needs to estimate the rf failure levels for a certain device, or the requirements of a Zener protection circuit for a VLSI circuit. Did the junction or transistor itself fail, or is the device shunted by a shorted path caused by a filamentary breakdown at the aforementioned interface, or both?

When failures occur only at this interface, improving either the junction topology and size or the metallic contact alone will not necessarily reduce the susceptibility of the device to rf transient burnout. Although efforts can be made to avoid sharp corners on the surface junction topology, it is also necessary to consider the effects of

Figure 4 . (a) Optical image and (c) x-ray topographic image recorded in grazing Bragg-Laue geometry following high-power-microwave-induced breakdown. Arrows point to damaging filaments. Image (b) is an x-ray topographic image for an unfailed triangular junction.



similar topologies from the cross-sectional point of view (which may be easily forgotten since this view is not readily observable).

In light of our results, it would not be surprising that two equal devices, one being a planar device and the other a vertically structured device (i.e., a 1N4148 silicon diode and equivalent size planar device), would fail at different rf power levels. This suggests important limitations on any rf burnout model or code, and that rf failure data for planar integrated devices in component reliability databases may be suspect.

Another important observation made when the x-ray results were reviewed is that surface evidence of failure, as seen optically, may not correspond with the

extent and location of crystal damage. Some devices actually failed that showed little or no surface evidence, yet had catastrophic crystal damage that only x-ray topographic imaging could reveal. Figures 2 and 3 show the comparison of optical micrograph and x-ray synchrotron images of the rf-induced failure in a 300- μm -diameter silicon planar *pn* junction. Note the characteristic contrast feature associated with the filament.

In figure 2(a) one can observe three spot-like features on the metallization: one not too far from the center of the circular anode, a second one in the square contact pad (this corresponds to the position of the probe contact), and a third one at the metal junction interface, as shown by the arrow. Of the two spots observed within the junction region, the one near the center is only a surface defect on the metallization, while the other (indicated by the arrow) is the visible portion of a filamentation channel which has a region of severe lattice damage associated with it just below the surface. Employing the grazing Bragg-Laue geometry, with its variable penetration depth, enables one to conclude that the maximum lattice distortion occurs closer to the surface than to the internal interface between the substrate and the epitaxial region of the junction (see Dudley et al [4]).

Damage induced by HPM waveforms shows a similar pattern. Figure 4(a) reveals what appears to be a filament at the interface between the metal contact and the diode itself. Comparison between the x-ray topographic contrast features associated with this filamentary feature (on fig. 4(c)) and those observed from filaments induced in reverse bias reveals this similarity.

Although the synchrotron has not previously been used to probe active semiconductor devices, our research shows that the synchrotron is a unique and powerful diagnostic tool to image failures in semiconductor devices. This research tool is unique in spotting failures that are not visible to ordinary surface imaging methods. Based on these capabilities, it is now possible to better assess the susceptibility of different semiconductor technologies to rf-induced failures and understand how to improve the hardening of electronic systems. This work is continuing and is also probing new technologies such as silicon carbide (SiC) and GaAs, thus contributing to our understanding of the nature and causes of the damage mechanisms for the primary semiconductor technologies—an essential requirement for assessing the long-term survivability of battlefield assets employing electronics. This work thus is directly relevant to Army concerns for the survivability of future military systems employing semiconductor devices in high transient rf environments, such as HPM and EMP.

Acknowledgements

The support of the Harry Diamond Laboratories Semiconductor Engineering and Materials Technology Facility is acknowledged for the silicon wafer processing, without which this study would not have been possible.

Synchrotron topography was carried out at the National Synchrotron Light Source, which is funded by the Department of Energy.

References

- 1 L.W. Ricketts, J.E. Bridges, and J. Miletta, *EMP Radiation and Protective Techniques*, Wiley Interscience (1976), pp 78-79.
- 2 G.-D. Yao, M. Dudley, and J. Wu, *J. X-ray Science Tech.* **2**, 195 (1990).
- 3 M. Dudley, G. Tolis, D. Gordon-Smith and C. Fazi, Government Microcircuit Applications Conference Digest of Papers, Vol. XVI, 651 (1990).
- 4 C. Fazi, Ph.D. Thesis, Johns Hopkins University (1990).
- 5 M. Dudley, J. Wu and G.-D. Yao, *Nucl. Instr. & Meth.* **B40/41**, 388 (1989).
- 6 M. Dudley, J. Wu, G.-D. Yao, H.-Y. Liu, and Y.C. Kao, *Mat. Res. Symp. Proc.* **160**, 469 (1990).
- 7 M. Dudley, J. Wu, G.-D. Yao, and H.-Y. Liu, *Mat. Res. Symp. Proc.* **167**, 1031 (1990).
- 8 M. Dudley, G.-D. Yao, D. Paine, D. Howard, and R.N. Sacks, *Mater. Sci. & Engin.* **B10**, 75 (1991).

FISHER

Characterizing the Response of Solid Rocket Motors to Fragment Impact

Jamie M. Fisher, Mrs.*

Commander

US Army Missile Command

AMSMI-RD-PR-M

Redstone Arsenal, AL 35898-5249

I. Background

A survivability enhancement program for an Army air defense missile system was the origination of the work discussed here. In a battlefield environment survivability test series, the missile system reacted violently to a threat stimulus. The system sustained sufficient damage to represent a loss of both system and crew. Following this unacceptable performance, a survivability working group (SWG) was established to develop, investigate, and demonstrate potential survivability enhancements. The solid propellant rocket motor and high explosive (HE) warhead were the missile components which posed the greatest risk to the system and crew survivability. The rocket motor by virtue of its larger size, correspondingly larger volume of energetic material and the more sensitive nature of its composition, was considered the most vulnerable missile component.

The rocket motor consists of a filament wound, S-glass composite case with a nominal wall thickness of 0.089 inch. The case is lined and insulated prior to propellant casting. The liner is a two part system of neoprene and polyvinyl. The insulation, a styrene-butadiene rubber (SBR) has a thickness of 0.015 inch. The propellant is a hazards class 1.1, minimum smoke, cross-linked, double-base formulation containing cyclotrimethylene trinitramine (RDX). The rocket motor contains approximately 40 pounds of this detonable material.

State of the art technology did not offer an alternative to the detonable material which was capable of providing the required missile performance. Therefore, it became the focus of the working group to develop mitigative techniques involving external barriers to protect the energetic missile components from external threat stimuli. It was agreed that certain threats such as a shaped charge could not be defeated or mitigated significantly without completely redesigning the missile system. As redesign was an unacceptable solution, artillery was selected as a representative, yet defeatable threat stimuli.

In order to develop survivability improvements, it was necessary to define the level of protection required to prevent detonation of the missile's energetic components. By characterizing the sensitivity of the rocket motor and warhead, performance requirements for armoring concepts could be established. By armoring to realistic requirements, an optimized system providing maximum protection with minimal adverse impact on the system as it is currently designed could be achieved. In the past overly conservative and unrealistic protection requirements have hampered the development and implementation of armor and other survivability enhancement concepts.

As a member of the working group, the author was tasked to provide a characterization of the rocket motor's sensitivity to fragment impact for use in evaluating potential armor concepts. This characterization had to be accurate, timely and cost effective. Previously proposed test plans involved full scale motors subjected to fragment spray from detonating artillery shells. These tests are costly, time consuming, and often yield ambiguous results. In this instance, the cost of a full scale motor, off the production line, was \$14K. Artillery shells do not fragment in perfectly repeatable patterns, therefore every test is basically unique. Ten and more tests were suggested to obtain a good statistical data base. This would require a minimum of \$140K in rocket motor test assets alone.

Testing of this nature requires high speed instrumentation as well as manual data acquisition. Both increase test time and cost. High speed cameras are required to record the impact and any resulting reaction. The films must then be reviewed to ascertain the number of fragment impacts, location of impacts, fragment velocity at impact, fragment shape and orientation at impact as well as the time to reaction and nature of reaction. This can be a tedious, time consuming, subjective task. If the test article does not react, the impact sites must be located, counted, mapped, the depth of penetration measured, and if fragments are recovered, their geometry and mass recorded. From this collection of data, generalized statements are made relating the reaction of the test article to impact by various size/mass distributions of fragments from specific artillery shells.

II. Characterization Procedure

The author proposed an innovative, more efficient alternative characterization procedure. This proposed procedure involved the application of experimental and analytical techniques developed for explosives. While these tools have been widely used in the explosives industry, they have not been applied to Army solid propellant rocket motors. The proposed method was a four step process. The first step was to obtain equation of

state data on the solid propellant involved. In the second step, this data is used to conduct hydrocode simulations of fragment impact on the rocket motor. The simulation results can then be used to design a test plan tailored to provide a characterization of the detonation threshold for several fragment masses/geometries. Finally, in the fourth step, fragment impact tests are conducted to verify simulation results, further quantify critical mass/velocity combinations which result in motor detonations and provide a required level of protection to evaluate armor candidates against. This characterization procedure was completed within 4 months at an estimated total cost of less than \$150K.

A. Step One: Propellant Sensitivity Testing

The first step was to experimentally characterize the shock sensitivity of the solid propellant so that it could be numerically modeled. This was done by conducting a series of wedge tests. The specimen configuration, shown in Figure 1, is milled from propellant slabs as illustrated in Figure 2. The propellant wedge is then subjected to a planar shock wave by a donor charge detonation. A streak camera is used to monitor the progression of this wave up the angled face of the wedge. Donor charges are designed to introduce shock pressures of specific magnitudes. If the pressure is sufficient, the shock will transition to detonation. The streak film records both the distance the shock wave traveled before transitioning, and the speed of the wave. A series of tests were conducted over a range of input pressures. The input pressures are plotted against the associated run distances to yield a Pop plot. The shock Hugoniot, shock velocity-vs-particle velocity, is also obtained from wedge test data. The results for the solid propellant tested in this program are shown

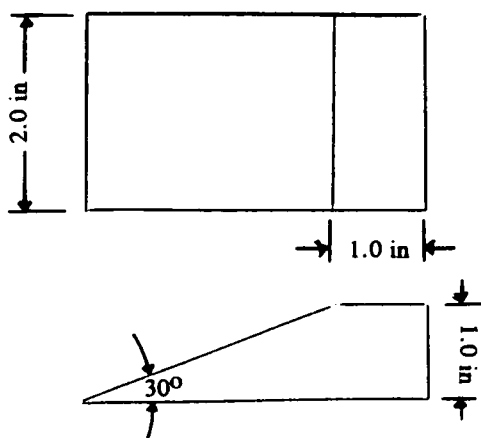


Figure 1. Wedge Specimen Configuration



Figure 2. Wedge Specimen Preparation

in Figures 3 and 4. These tests were conducted by Mr. Frederick Sandstrom at the Center for Explosives Technology Research of the New Mexico Institute of Mining and Technology in Socorro, New Mexico.

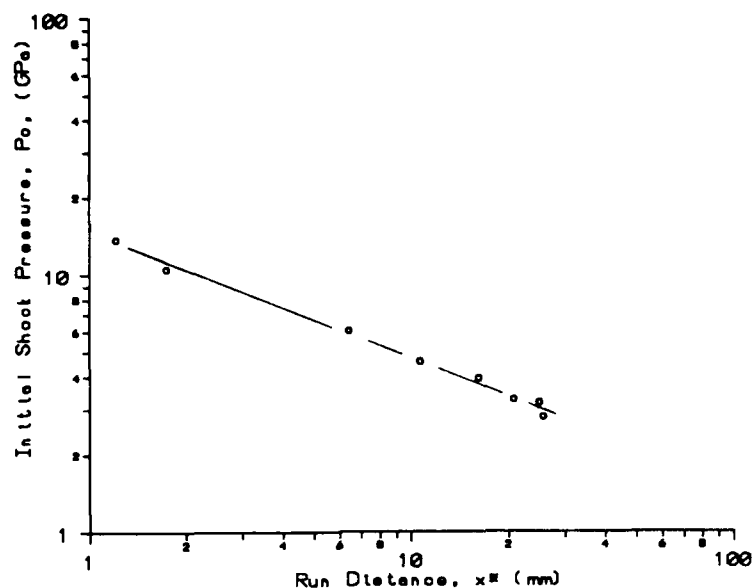


Figure 3. Propellant Pop Plot: Pressure-vs-Run Distance

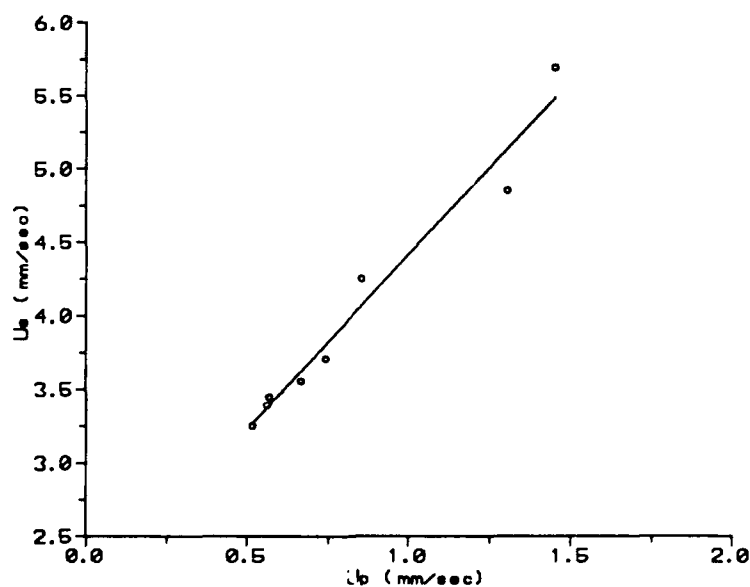


Figure 4. Propellant Shock Hugoniot: Shock Velocity-vs-Particle Velocity

B. Step Two: Hydrocode Simulations

Experimentally determined Pop plot and Hugoniot relations serve as inputs into a hydrocode simulation of impact scenarios. The hydrocode used in this exercise, 2DE, was developed at Los Alamos National Laboratory.¹ It is a Eulerian hydrocode which has been interfaced with an initiation code known as Forest Fire. Early attempts to model the propagation of shock waves treated all materials as fluids by not including strength effects. The term, hydrocodes, is a carry over from these early computer codes.

2DE, as are all hydrocodes, is based on classical continuum mechanics, that is conservation of mass, momentum, and energy. It is a two-dimensional, finite element (FE) code which uses a spatial representation of the problem. As such, the FE grid remains fixed throughout the calculations while material, momentum and energy are allowed to flow across cell boundaries. As time progresses, the code computes velocity, pressure, mass, density, etc., by applying the Eulerian form of the conservation equations as quantities of energy, mass and/or momentum flow in and out of the cells.²

In applying hydrocodes to explosives, additional input data is required to relate the explosive's material properties, density and internal energy, to pressure. Therefore, an equation of state for both unreacted and reacted material is required. The Forest Fire model utilizes the equations of state and the Pop plot, from wedge test data, to determine the reaction of an impacted material and the rate of explosive decomposition as a function of impact pressure. Forest Fire was written to predict the initiation and propagation of detonations in heterogeneous explosives. When an explosive is impacted with sufficient velocity, the induced shock pressure results in the decomposition of the explosive. This decomposition adds energy to the shock and detonation occurs.³

The numerical simulations conducted for this program modeled the impact of a fragment simulating projectile (FSP) on the rocket motor. The FSP was chosen over a 1/2 inch cube fragment because of the standard use of FSPs in armor evaluation. Candidate armors were to be tested against fragment mass/velocity combinations determined from the rocket motor testing. Consistency of the fragment geometry between the two experimental programs was therefore desirable. The FSP is also designed to be shot with standard powder guns and barrels. This simplifies the launching procedure for the experimental programs. The fragment geometry, material, and manufacturing are specified in MIL-P-46593A (Ord).

Three FSP sizes were chosen to obtain detonation threshold velocities for the rocket motor across a band of fragment masses. The three FSPs used were a 44 grain (or caliber 30), a 207 grain (or caliber 50), and a 830 grain (or 20 mm). Figure 5 illustrates the geometry of the 207 grain FSP which is then scaled for the other fragment masses. All three FSPs are shown in Figure 6. The 2DE simulations were conducted by Dr. Tom Cost, Athena Engineering Company, and yielded the following detonation threshold velocities for the rocket motor. The 44 grain fragment required a velocity of 7743 ft/sec to detonate the motor. The 207 grain required 4255 ft/sec, and the 830 grain required 2502 ft/sec.

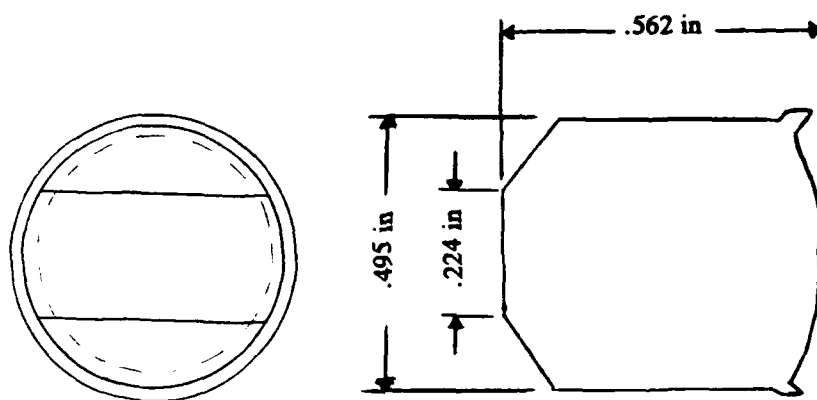


Figure 5. Fragment Simulating Projectile Configuration

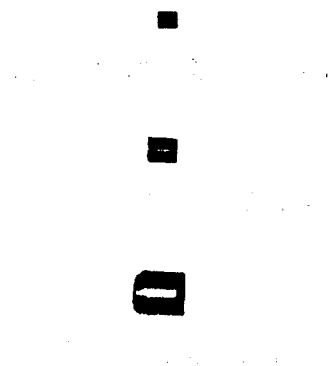


Figure 6. FSPs used in Motor Impact Testing

C. Step Three-Test Plan Development

From previous experience, it was known that a test article with a L/D of one or greater was sufficient to yield the same type of response to impact as a full scale motor. It was therefore recommended to dissect full scale motors into thirds. This reduced test costs and provided more test specimens. Non-deliverable rocket motors with discrepancies in the position of the wiring harness were available from the motor manufacturer. These were purchased at a reduced price. The discrepancy in the wiring harness would not affect the motor response to fragment impact. Therefore these motors were acceptable for this test program. The purchase of these motors resulted in additional program cost savings.

The motors were dissected using a remotely operated, high-pressure grit blast technique. Each motor section was then fitted with two, 1-inch thick, aluminum end plates held by four steel tie rods. The sections were potted in place. One plate contained a one inch dia. hole to serve as a nozzle, providing venting while maintaining internal pressures sufficient to sustain a reaction. The test vessel configuration is shown in Figure 7.

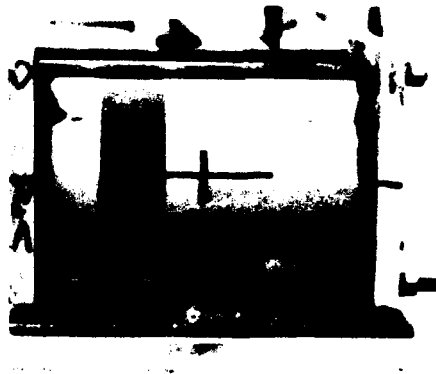


Figure 7. Rocket Motor Section Test Article

Vulnerability assessments conducted for this missile system were based on the assumption that any fragment with a velocity of 1250 ft/sec or higher would cause the rocket motor to detonate. Based on experience, previous tests with this system and the hydrocode results, this assumption was believed to be overly conservative. Therefore, it was not used as the starting velocity for the experimental motor impact program. Rather, the test plan to experimentally characterize the rocket motor's impact sensitivity was

designed to start with the 207 grain fragment at a velocity of 4000 ± 20 ft/sec. The following shots were planned at 300ft/sec increments until the response went from no detonation to detonation, or vice versa. Once the response change was achieved, the next shots would be at a velocity halfway between those resulting in the two different reactions. This 'splitting the difference' proceeds until the detonation threshold velocity is determined to within a maximum of 50 ft/sec. This technique is the quickest way to zero in on the desired value, saving both time and test assets. The same procedure would be repeated for the 830 grain FSP. The starting velocity, based on numerical simulation results, would be 2500 ± 20 ft/sec.

The predicted detonation velocity for the 44 grain FSP is above the capability of the caliber 30 powder gun. Therefore, it was recommended that the starting velocity be the maximum velocity achievable with the standard caliber 30 gun hardware. Further testing would be based on the velocity achieved and the motor response at this velocity. It was expected that the predicted velocity for the 44 grain FSP could not be verified experimentally without switching to a larger barrel and saboting the fragment. The utility of such an exercise would be determined following the completion of all other FSP shots.

D. Step Four: Rocket Motor Impact Testing

The rocket motor fragment impact tests were conducted at the New Mexico Institute of Mining Technology in Socorro, NM. The test set up is shown in Figure 8. Two blast gages were placed 10 feet in front of the test article, one on each side, along 45° angles. Two high-speed cameras were used to record impact and motor response. In addition, two real time video recordings were taken during the entire test procedure. Fragment velocities were determined with a Doppler radar as well as multiple velocity screens. The break screens used to determine velocity, also provided information on the orientation of the fragment by the petalling pattern left around the hole as the fragment penetrated the screen. Table I contains a summary of the fragment impact velocities and associated motor response for the 207 grain FSP. Figures 9 through 11 show representative examples of the motor responses obtained during this fragment testing. Figure 9 shows a motor which was penetrated by the fragment, ignited and as a result burned quiescently. Figure 10 shows a motor section which was penetrated by a fragment with no subsequent reaction. Figure 11 shows the aftermath of a motor section which detonated upon impact. As expected, there were no recoverable materials from the motor section, witness plate, or test stand following a detonation.

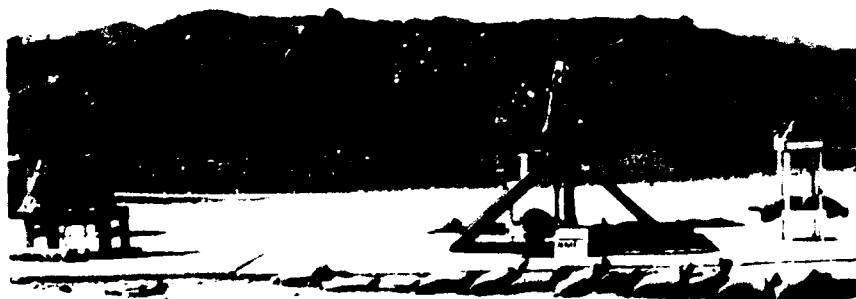


Figure 8. Fragment Impact Test Setup



Figure 9. Rocket Motor Burn Response to Fragment Impact



Figure 10. Rocket Motor with No Response to Fragment Impact



Figure 11. Rocket Motor Detonation Response to Fragment Impact

Table I. Test Results for 207 grain FSP

Impact Velocity	Reaction
3944	Burn
4016	No Reaction
4050	Burn
4083	Burn
4283	Detonation
4300	Burn
4326	Detonation
4547	Detonation

The detonation threshold velocity for the 207 grain FSP was determined to be 4313 ft/sec. The detonation response at the lower velocity, 4283 ft/sec, was due to fragment tumbling at impact. Difficulties in reaching the desired velocities were encountered with the FSP. The required shot velocities were well above the normal muzzle velocity of a caliber 50 gun, which the FSPs were designed for. The additional pressure required to achieve the higher velocities exceeded the shear strength of the fragment obturator. When the obturator failed, pressure was allowed to leak past the fragment, thus lowering the shot velocity. This problem was overcome by modifying the

fragments. The obturator was removed and a polypropylene pusher plug was attached to the rear of the fragment. This modification facilitated more predictable and repeatable shot velocities with only a slight drop in the total fragment mass. The modified fragment configuration is shown in Figure 12.

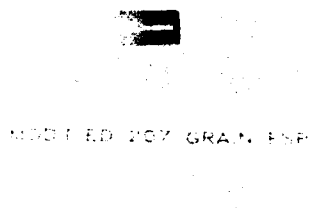


Figure 12. FSP Modified Configuration

Table II contains a summary of the impact velocities and associated motor response for the 830 grain FSP. Seven shots were required to determine the detonation threshold velocity of 1900 ft/sec. With this large fragment the motor response was limited to 'no response' or 'detonation'. The fragment impact did not induce burning at velocities below the threshold value as was seen with the 207 grain FSP. Examining the motor sections which were impacted with no response, showed no indication at the entry site or along the bore surface of localized ignition or burning.

Table II. Test Results for 830 grain FSP

Impact Velocity	Reaction
1767	No Reaction
1840	No Reaction
1853	No Reaction
1872	No Reaction
1928	Detonation
2036	Detonation
2700	Detonation

Only two shots were conducted with the 44 grain FSP. As discussed previously, the predicted detonation threshold velocity was too high to verify experimentally with a standard set up. The maximum explosive powder load for the caliber 30 gun was used, which yielded impact velocities of 4858 ft/sec and 4965 ft/sec. In both instances the

fragment failed to exit the motor section. The hot fragment embedded in the propellant web, causing local ignition which spread relatively slowly to the bore surface, igniting the rest of the motor. This is evidenced by the video which shows chuffing at the entry hole for several seconds, then flames appearing in the entry hole, followed shortly by smoke and flames exiting the simulated nozzle. Both motor sections burned for approximately 6 minutes, which was three times the length of burn experienced with the 207 grain fragment. Due to the benign reaction of these two shots and the good correlation between predicted and experimental results for the 207 and 830 grain FSPs, it was decided to suspend further testing with the 44 grain fragment. It was also noted that the predicted detonation threshold velocity, 7743 ft/sec, was well above the fragment velocity from artillery threats of interest to this program.

III. Application of Characterization Results

The rocket motor impact test program served three major purposes. First, it provided a timely, accurate, and cost effective characterization of the motor's sensitivity to fragment impact. This data was used to correct and update vulnerability analyses for the missile system. The previously discussed detonation velocity of 1250 ft/sec was proven to be overly conservative and invalid for survivability enhancement studies.

Secondly, the experimental program validated the results of the numerical simulations. Figure 13 shows the predicted detonation threshold velocities plotted as a function of fragment mass. Also shown are the experimentally determined values. For the 207 grain FSP, the predicted value was 1.3 percent lower than the experimental result. For the 830 grain FSP, the predicted value was 24 percent higher than the experimentally determined velocity. These are good correlations between analytical and experimental results of this nature. The results from the hydrocode calculations were certainly more accurate than the detonation threshold velocity value used prior to this test program.

The third and primary goal of the rocket motor impact test series was to provide a criteria against which to evaluate potential armor solutions. The test program was successful in meeting this goal. The candidate armors were tested with the 207 grain FSP at a velocity of 5000 ft/sec. This fragment mass/velocity combination was above the threshold value for the rocket motor and provided a measure of the effectiveness of the armors at reducing the mass/velocity to below the critical value. Armors which were successful at reducing the mass/velocity to values below the detonation threshold offered potential survivability enhancement for the system.

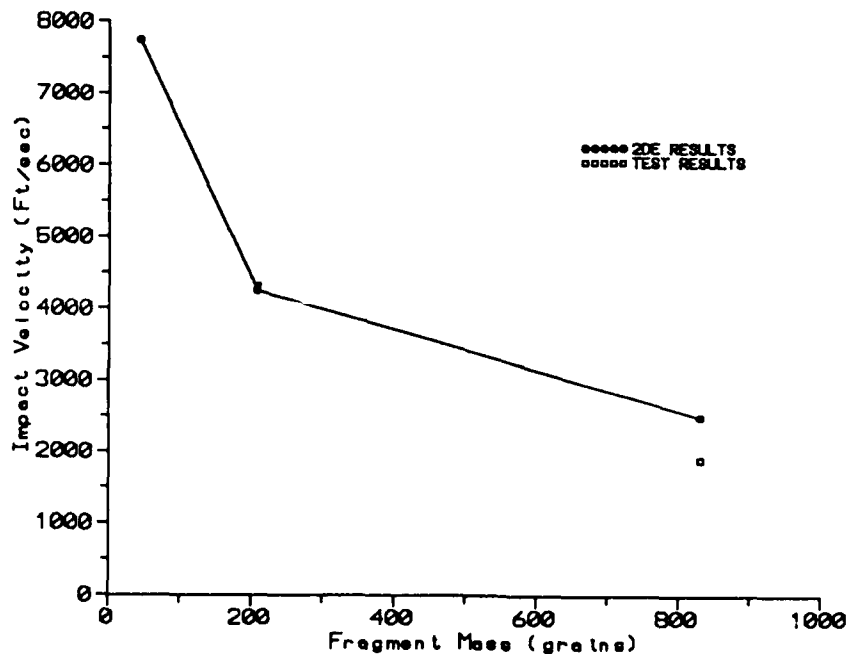


Figure 13. Rocket Motor Detonation Threshold Velocity:
Critical Velocity-vs-Fragment Mass

IV. Conclusions

As part of the survivability enhancement program for an Army missile system, an innovative methodology for characterizing the response of a solid propellant rocket motor to fragment impact was demonstrated. This methodology combined experimental techniques developed for high explosives with an analytical tool derived from basic continuum mechanics.

Wedge tests were conducted on the solid propellant to evaluate the shock sensitivity of the material and to derive its equation of state. The experimentally determined parameters were then input into 2DE, a two-dimensional hydrocode, to conduct numerical simulations of impacts against the rocket motor. The hydrocode analyses were used to predict the impact velocity required for a specific projectile configuration to initiate a detonation in the

rocket motor. With these predicted velocities, a test plan was tailored to experimentally characterize the detonation threshold for the rocket motor and significantly reduce the number of tests and assets required. In this program, the experimentally determined detonation velocities correlated well with the predicted values thereby offering validity to the application of these procedures.

Application of this methodology offers numerous advantages. As demonstrated in this program, this procedure is a viable alternative to the time consuming and costly experimental programs normally used to characterize the response of a solid rocket motor to fragment impact. It also provides a procedure to screen developmental propellants, in terms of their shock sensitivity, prior to committing them to significant production scale up. The combination of testing and modeling utilized here also provides a more accurate data base for designers to use in investigating mitigative techniques such as shields or armors to enhance system survivability. In addition, the procedure demonstrated here offers a tool for the evaluation of fragment and bullet impact response of solid rocket motors for Insensitive Munitions (IM) considerations. This procedure offers to IM investigators, as it did to the SWG, a more cost-effective and timely means of characterizing the response of a solid propellant rocket motor.

V. Future Plans

The success of this innovative methodology which was implemented as part of a survivability enhancement program has sparked interest in its use on other Army systems currently involved in similar activities and as a tool for IM modeling. To further investigate, validate, and improve the effectiveness of this characterization procedure, several follow-on efforts are being planned. The first will include additional modeling and testing of the currently characterized solid propellant against different fragment geometries. Characterization of the shock sensitivity, through wedge testing, of additional Army propellants as well as developmental IM type propellants is underway. Other hydrocodes, both one and two-dimensional, will be evaluated as to their modeling and predictive capabilities. The most effective and useful codes will then be brought in-house to provide researchers with an accurate modeling tool for impact scenarios.

FISHER

REFERENCES

1. J. D. Kershner and C.L. Mader, *2DE: A Two-Dimensional, Continuous, Eulerian Hydrodynamic Code for Computing Multicomponent Reactive Hydrodynamic Problems*, Los Alamos Scientific Laboratory, LA-4846, March 1972.
2. James Wilbeck, Charles Anderson, et al., *A Short Course on Penetration Mechanics*, Southwest Research Institute, San Antonio, TX, undated.
3. Charles L. Mader and Charles A. Forest, *Two-Dimension Homogeneous and Heterogeneous Detonation Wave Propagation*, Los Alamos Scientific Laboratory, LA-6259, June 1976.

FRIEDLANDER, WELKOS, PITT, EZZELL, WORSHAM, ROSE, IVINS

TITLE: Post-exposure Prophylaxis against Experimental Inhalation Anthrax (U)

*Arthur M. Friedlander, COL, Susan L. Welkos, Dr.,
Margaret L. M. Pitt, Dr., John W. Ezzell, Dr.,
Patricia L. Worsham, Dr., Kenneth J. Rose, SFC,
Bruce E. Ivins, Dr., John R. Lowe, LTC,
Gerald B. Howe, SFC, Perry Mikesell, MAJ, and
Wade B. Lawrence, MAJ.

U. S. Army Medical Research Institute of Infectious
Diseases, Fort Detrick, Frederick, Maryland 21702-5011.

ABSTRACT: Inhalation anthrax is a rare disease that is almost invariably fatal. In this study, we determined whether a prolonged course of post-exposure antibiotics with or without vaccination would protect when the drug was discontinued. Rhesus monkeys were exposed to a lethal aerosol dose of *Bacillus anthracis*. Beginning 1 day after exposure, groups of 10 animals were given penicillin, ciprofloxacin, doxycycline, doxycycline plus vaccination, vaccination alone, or saline. Antibiotics were administered for 30 days and then discontinued. Vaccine was given on days 1 and 15 postexposure. Two animals died of causes other than anthrax and were not included in the statistical analysis. Nine of 10 controls and 8/10 animals given vaccine alone died. There were no deaths due to anthrax in any antibiotic-treated animal while on therapy. Each antibiotic regimen provided significant protection upon discontinuance of the drug (penicillin, 7/10 survived, $P < 0.02$; ciprofloxacin, 8/9 survived, $P < 0.002$; doxycycline, 9/10 survived, $P < 0.002$; doxycycline plus vaccination, 9/9 survived, $P < 0.0002$). Thus, post-exposure treatment of inhalation anthrax with antibiotics for 30 days beginning 1 day after exposure completely protected animals while on treatment and provided significant long-term protection upon discontinuance of therapy. Protection against re-challenge was provided by combining post-exposure antibiotic treatment with vaccination.

BIOGRAPHY: Arthur M. Friedlander, COL

PRESENT ASSIGNMENT: Chief, Bacteriology Division, U.S.
Army Medical Research Institute of Infectious Diseases, Fort Detrick,
Frederick, Maryland 21702-5011.

DEGREES HELD: A.B., Harvard College, MA, 1961; M.D.,
University of Pittsburgh, PA, 1965.

Post-exposure Prophylaxis against Experimental Inhalation Anthrax (U)

Arthur M. Friedlander, COL*, Susan L. Welkos, Dr.,
Margaret L. M. Pitt, Dr., John W. Ezzell, Dr., Patricia L. Worsham, Dr.,
Kenneth J. Rose, SFC, Bruce E. Ivins, Dr., John R. Lowe, LTC,
Gerald B. Howe, SFC, Perry Mikesell, MAJ and Wade B. Lawrence MAJ
U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick,
Frederick, Maryland 21702-5011.

Anthrax is a zoonotic infection, mainly of herbivores, caused by *Bacillus anthracis*. Humans become infected by contact with infected animals or contaminated animal products. Anthrax begins by introduction of the spore through: 1) the skin producing cutaneous anthrax, 2) the gastrointestinal tract causing oropharyngeal or gastrointestinal anthrax, or 3) the respiratory tract causing inhalation or mediastinal anthrax. The inhalation form of the disease was first recognized in the 19th century during the Industrial Revolution with the development of the textile and tanning industries and was one of the first occupational pulmonary diseases identified¹. Workers in these industries were exposed to high concentrations of anthrax spores in the air, which were generated by the processing of contaminated wool, hair, or hides². Presently, inhalation anthrax is extremely rare, with approximately 30 cases reported in this century, most often associated with industrial exposure to spores^{3,4}. Inhalation anthrax has been almost uniformly fatal because of the difficulty in establishing the diagnosis and the rapid progression of the disease. Previous experimental studies demonstrated that treatment with penicillin for 5 or 10 days, beginning 1 day after aerosol exposure of rhesus monkeys, was protective during the drug therapy, but animals died when the antibiotic was discontinued⁵. Long-term protection was afforded only by combining penicillin therapy with post-exposure vaccination. Recent events in the Gulf War heightened the awareness of the possibility of using anthrax as a biological weapon. For this reason, we determined whether a more

prolonged course of antibiotic therapy with or without vaccination, begun after an aerosol exposure, could protect animals from lethal inhalation anthrax. We also determined if any long-term survivors were immune upon subsequent re-challenge.

Materials and Methods

Bacterial strain preparation, animals, and aerosol exposure.

Bacillus anthracis Vollum 1B spores were prepared and stored as previously described, except for the omission of centrifugation through Renografin gradients⁶. On the day of the experiment, spores were diluted in sterile water, heated at 60°C for 45 min, and then kept on ice. Rhesus monkeys (*Macaca mulatta*) of both sexes, weighing from 5.8 to 13 kg, were anesthetized with tiletamine/zolazepam (3 mg/kg, A. H. Robins Co., Richmond, VA) and exposed in a head-only 1 ft.³ plexiglass chamber contained within a Class III biological safety cabinet to an aerosol generated with a Collison nebulizer for 10 min followed by a 5 min air wash. The mass median diameter of the particles generated was 1.2µm as determined by cascade impactor and aerodynamic particle sizer (Model 3310, TSI Incorporated, St. Paul, MN). The concentration of bacterial spores in the aerosol was determined during exposure for each monkey by sampling the aerosol with an all-glass impinger. The minute respiratory volume was measured on each anesthetized animal immediately before exposure. In the first challenge experiment, animals were exposed to an inhaled dose of $4.0 \pm 1.6 \times 10^5$ spores (mean \pm SD), corresponding to approximately 8 LD₅₀ [5, Welkos, S. L. et al., unpublished data]. All survivors from the first experiment were re-challenged by aerosol with an inhaled dose of $2.6 \pm 1.4 \times 10^6$ spores corresponding to approximately 50 LD₅₀.

Experimental groups. The animals were distributed randomly by sex and weight in six experimental groups. 1) Controls: 10 animals were given saline as a control solution intramuscularly every 12 hours, beginning 1 day after exposure. 2) Penicillin: 10 animals were given procaine penicillin G intramuscularly at a dose of 180,000 units (0.6 ml) every 12 hours, beginning 1 day after exposure and continuing for 30 days. 3) Ciprofloxacin: 10 animals were given ciprofloxacin at a dose of 125 mg (in 5 ml H₂O) by orogastric tube every 12 hours, beginning 1 day after exposure and continuing for 30 days. 4) Doxycycline: 10 animals were given doxycycline at a dose of 30 mg (6 ml) by orogastric tube every 12 hours,

beginning 1 day after exposure and continuing for 30 days. 5) Doxycycline plus vaccination: 10 animals were given doxycycline at a dose of 30 mg (6 ml) by orogastric tube every 12 hours, beginning 1 day after exposure and continuing for 30 days. They were also given 0.5 ml of the human anthrax vaccine (Michigan Department of Public Health, lot #FAV001) on days 1 and 15 after the aerosol exposure. 6) Vaccination: 10 animals were given 0.5 ml of the human anthrax vaccine on day 1 and, if still alive, on day 15 after the aerosol exposure. They were also given water by orogastric tube every 12 hours, beginning 1 day after exposure and continuing for 30 days. Animals given ciprofloxacin, doxycycline, or vaccination were anesthetized with tiletamine/zolazepam (3 mg/kg) every 12 hours in order to give the medication by orogastric tube. After 30 days of treatment, antibiotics were discontinued and the animals were given no additional therapy. Survivors were re-challenged 131 to 142 days after the initial exposure together with five additional new untreated controls.

Clinical, microbiological, and pathological studies. Daily blood cultures were obtained from the untreated controls and vaccination groups until death or for 14 days. In the antibiotic-treated groups, blood was cultured every other day until 80% of the controls died, then twice weekly until day 30 when antibiotics were discontinued, then every other day until approximately day 60, and then once a week until re-challenge. Blood from untreated animals was collected in an Isolator 1.5 (Du Pont Co., Wilmington, DE) and cultured in 10-fold dilutions in triplicate on trypticase soy agar. Blood from treated animals collected in an Isolator 1.5 was cultured undiluted and at a 1:100 dilution on trypticase soy agar. In addition, 1 ml was cultured in a Bactec Peds Plus bottle (Becton Dickinson, Towson, MD). Blood obtained before and at various times after challenge was analyzed for IgG antibodies to the anthrax protective antigen by an ELISA with protein A-horse radish peroxidase (Boehringer Mannheim, Indianapolis, IN) and 2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid) as substrate (Sigma, St. Louis, MO). Blood was also analyzed at various times for routine hematology and biochemistry.

All animals were observed at least twice daily until death. Moribund animals were killed by deep anesthesia (tiletamine/zolazepam, 6 mg/kg) and exsanguination. All animals were autopsied. A diagnosis of anthrax was confirmed in all animals by isolating *B. anthracis* from the blood. Bacteria were identified by typical morphology, production of capsule when grown in

20% CO₂, and production of protective antigen as detected by a precipitin halo when they were grown on agar plates containing antibody to protective antigen. In some cases, organs were cultured quantitatively. In all deaths in which antemortem blood cultures were negative, cultures were obtained at autopsy of the blood, spleen, lung, liver, intrathoracic lymph nodes, and brain.

Antibiotic sensitivity testing and serum levels. Minimal inhibitory concentrations (MIC) of the *B. anthracis* Vollum 1B strain were determined in Mueller-Hinton broth dilutions by using an inoculum of $2.5-3 \times 10^5$ /ml in tubes and in a microtiter format. The MIC was 0.08 µg/ml for penicillin, 0.08 µg/ml for ciprofloxacin, and 0.02 µg/ml for doxycycline. The minimal bactericidal concentration was 0.32 µg/ml for penicillin and 0.08 µg/ml for ciprofloxacin.

Peak levels of antibiotics in the serum were determined 1 hour (ciprofloxacin) or 2 hours (penicillin and doxycycline) after a dose. Trough levels were measured 12 hours after a dose for all drugs.

Procaine penicillin G, penicillin G potassium used as a reference standard, and doxycycline monohydrate suspension were purchased from Pfizer, Inc. (New York, NY). Powdered doxycycline hyclate for a reference standard was a gift from Pfizer, Inc. (Groton, CT). Ciprofloxacin tablets and powdered ciprofloxacin for a reference standard were gifts of Miles Pharmaceuticals, Inc. (West Haven, CT).

The experiments were carried out under the guidance of the Veterinary Medicine Division in compliance with the Animal Welfare Act and other Federal statutes and regulations relating to animals and experiments involving animals and adhere to principles stated in the Guide for the Care and Use of Laboratory Animals, NIH publication 86-23, 1985 edition. The facilities are fully accredited by the American Association for Accreditation of Laboratory Animal Care.

Results

Description of the disease in untreated controls. Nine of the 10 untreated control animals exposed to an inhaled dose of 8 LD₅₀ died within 3 to 8 days following the challenge (mean \pm SE = 5.6 \pm 1.1 days). The animals were ill for 1 to 4 days before death, demonstrating decreased spontaneous activity, weakness, and anorexia. One animal had a single

seizure on the day of death and was found to have meningitis on autopsy. Respiratory distress was observed in only one animal. Bacteremia at levels of 10 to 10^5 colony forming units (CFU)/ml, was present for a mean of 1.8 ± 0.9 days before death. Terminal bacteremias in eight of the nine animals that died varied from 10^4 to 10^9 CFU/ml. The one animal with a low terminal bacteremia of 2×10^2 CFU/ml had meningitis with 2×10^7 CFU/g of brain tissue. Five of nine animals had gross findings of mediastinitis and enlarged intrathoracic nodes with hemorrhagic lymphadenitis. A glistening edema in the mediastinum was also present in several animals. Meningitis was present in five of nine animals and was hemorrhagic in three of the cases. The one animal that survived never had a positive blood culture.

Antibiotic serum levels. The mean peak and trough serum antibiotic levels for each of the antibiotics did not vary significantly when measured on days 5, 9, 20, and 30. The peak levels were at least 10 times the MIC for all antibiotics and the trough levels varied from 1 times the MIC for ciprofloxacin to 1-10 times for penicillin and doxycycline (Kelly, D. et al., manuscript submitted).

Effect of post-exposure treatment on survival. Survival of the various treatment groups is shown in Table 1. Eight of 10 animals treated with vaccination alone died 5 to 10 days after challenge (mean of 7.0 ± 1.5 days); which did not differ significantly from the untreated controls. The disease in the group given only vaccine was similar to that in the untreated controls. Those given vaccine alone were bacteremic for a mean of 1.9 ± 0.7 days before death at the same low levels observed in the controls. The degree of terminal bacteremia and the findings on autopsy were also similar to those observed in the untreated controls. Of the two vaccinated animals that survived, one had persistently negative blood cultures while the other had transiently positive blood cultures on days 5, 11, and 12 at low levels of 10-20 CFU/ml. Blood cultures in this animal remained persistently negative after day 12.

We observed significant protection against death in each of the antibiotic-treated groups. All animals in the penicillin group survived the 30 days while on treatment during which their blood cultures remained negative. Upon discontinuance of the penicillin, 3 of 10 animals died of anthrax. Deaths occurred on days 9, 12, and 20 after stopping the penicillin. In all three cases, death was due to anthrax.

Table 1. Survival after post-exposure treatment of inhalation anthrax

	<u>Anthrax deaths</u>	<u>P value (vs control)</u>
Control untreated	9/10	
Vaccine alone	8/10	>0.1
Penicillin	3/10	<0.02
Ciprofloxacin	1/9*	<0.002
Doxycycline	1/10	<0.002
Doxycycline+vaccine	0/9**	<0.0002

*One animal died 5 days after exposure from aspiration pneumonia and had no evidence of anthrax on autopsy. This animal was excluded from analysis. A second animal died due to urethral obstruction 73 days after discontinuance of ciprofloxacin and had no evidence of anthrax at autopsy. This animal is included in the statistical analysis as a survivor.

**One animal died 6 days after discontinuing doxycycline with no evidence of anthrax on autopsy. The cause of death remains unknown and the animal was excluded from statistical analysis.

One of the ciprofloxacin animals died 5 days after exposure to anthrax from an aspiration pneumonia that developed 24 hours after the inadvertent introduction of drug into the trachea. All antemortem blood cultures were negative for anthrax as were the post-mortem cultures of blood, lung, liver, spleen, and brain. Pathological analysis showed an aspiration pneumonia with no histological evidence of anthrax in any organ. This animal was eliminated from analysis as no assessment of the effectiveness of antibiotic treatment on long-term survival was possible. All other ciprofloxacin-treated animals survived the 30 days of treatment with negative blood cultures. One animal died of anthrax with a positive terminal blood culture

6 days after stopping the antibiotic. This animal also had severe endometriosis on autopsy. Another animal in the ciprofloxacin group died 73 days after discontinuance of antibiotic. Autopsy cultures of blood and organs were negative for anthrax and there was no other pathological evidence of anthrax. This animal died of urethral obstruction due to rubbery plugs (concretions) in the proximal urethra and bladder. We considered this animal to be a survivor of the anthrax challenge.

In the group treated with doxycycline alone, all animals survived during therapy and had negative blood cultures. One animal died of anthrax 28 days after treatment was stopped.

None of the animals treated with doxycycline plus post-exposure vaccination died of anthrax. One animal in this group died 6 days after discontinuance of doxycycline, but had no evidence of anthrax on autopsy by either culture or histologically. Mild myocardial degeneration and necrosis were observed, but the exact cause of death could not be definitively determined. The effect of treatment on long-term survival from anthrax could not be evaluated in this animal and it was eliminated from statistical analysis.

The gross and microscopic pathological findings in the antibiotic-treated animals that died of anthrax after discontinuance of the drug did not differ from those observed in the untreated controls.

The animals that survived exposure to the aerosol challenge were examined for evidence of an immune response. Sera obtained 131-142 days after exposure were tested for antibody to the protective antigen component of the anthrax toxin. As seen in Table 2, none of the surviving animals treated with penicillin, ciprofloxacin, or doxycycline alone had a demonstrable immune response. In contrast, the exposed and surviving animals that had been given vaccine in addition to doxycycline therapy all developed a fourfold or greater rise in antibody to protective antigen. In addition, the two surviving animals that had been given vaccine alone also developed an antibody response to protective antigen. Thus, the only anthrax-challenged animals to develop an immune response were those that were vaccinated.

Table 2. Serological response to anthrax protective antigen in survivors after aerosol challenge with anthrax spores

Group	≥Fourfold rise in antibody titer*
Control untreated	0/1**
Vaccine alone	2/2
Penicillin	0/6
Ciprofloxacin	0/7
Doxycycline	0/9
Doxycycline + vaccine	9/9
*Sera obtained 131-142 days after challenge were tested for antibody to anthrax protective antigen by ELISA as described in Materials and Methods.	
**Number of animals positive/number of animals tested.	

Resistance of surviving treated animals to re-challenge. Surviving animals were re-challenged with anthrax spores 131-142 days after the initial challenge. The dose was approximately 50 LD₅₀. As seen in Table 3, significant protection against re-challenge occurred only in the group that had been vaccinated and treated with doxycycline. There was no significant protection afforded by antibiotic treatment alone. Not shown in Table 3 is that the two animals that were given vaccine alone without doxycycline and that survived the initial challenge also survived the re-challenge. Thus, the only animals to survive re-challenge were those that had been vaccinated. These surviving animals have now been observed for 1 year after the re-challenge and have remained free of disease.

Table 3. Resistance of survivors of post-exposure treatment to re-challenge

Group	Anthrax deaths	P value (vs control)
Control*	4/5	
Penicillin	7/7	>0.1
Ciprofloxacin	6/7	>0.1
Doxycycline	9/9	>0.1
Doxycycline + vaccine	0/9	<0.005

*Controls consisted of five additional animals not previously exposed to anthrax.

Discussion

The clinical and pathological findings, including the duration of bacteremia preceding death, which we observed in rhesus monkeys after aerosol exposure to anthrax spores are consistent with previously reported studies^{7,8}. The observations are strikingly similar to those reported in cases of human inhalation anthrax,^{3,9,10} which confirms the relevance of this experimental model. The most striking pathological findings were the intrathoracic hemorrhagic lymphadenopathy, mediastinitis, and meningitis.

One of the control animals survived the initial aerosol challenge and appears not to have become infected as all blood cultures were negative and an immune response to protective antigen never developed (Table 2). There were two animals in the group given vaccine alone that also survived. One of these had a positive blood culture with low level bacteremia. This suggests that post-exposure vaccination alone may occasionally be protective. It is perhaps more likely that, on occasion, anthrax may not always be fatal in non-human primates. This is supported by the previous observations that

two of four chimpanzees exposed to an aerosol of anthrax spores had transient low-grade bacteremias and survived¹¹.

There were three deaths which were not due to anthrax. One of the ciprofloxacin-treated animals died of an aspiration pneumonia 5 days after starting treatment and had no evidence of anthrax at autopsy. One animal in the doxycycline plus vaccine group died 6 days after discontinuing the antibiotic on day 36. There was no evidence of anthrax on autopsy and the cause of death could not be established. These two animals were removed from statistical analysis as the effect of the treatment regimen could not be adequately evaluated. The third animal was in the ciprofloxacin group and died 73 days after discontinuance of the antibiotic without any evidence of anthrax on autopsy. This animal died of urethral obstruction due to rubbery concretions. The relationship of this finding to the crystalluria induced by ciprofloxacin in non-human primates is unclear as the animal had not been taking ciprofloxacin for more than 2 months. We considered this animal to be a survivor of anthrax.

Critical to the rational treatment of inhalation anthrax is an understanding of the initial pathogenesis of the disease. The spore is the naturally infectious form of the organism. Studies by Ross¹² suggest that inhaled spores are rapidly phagocytosed by alveolar macrophages and transported by draining lymphatics to the regional lymph nodes where germination to the vegetative bacillus occurs. With release of bacilli from phagocytes, there is rapid extracellular bacterial proliferation throughout the node with subsequent spread locally and to other organs. However, some of the inhaled spores remain dormant and do not germinate. Most important, from the perspective of therapy, these spores may remain within the lung for extended periods of time. Henderson et al. demonstrated that 42 days after inhalation of spores by rhesus monkeys, 15-20% of the initially retained dose of spores were still present in the lung⁵. The significance of this for therapy of anthrax was initially appreciated by Barnes as early as 1947, who stated in reference to penicillin that spores may persist in the tissues and germinate after the blood level falls¹³. The previous findings of Henderson et al. on post-exposure treatment of inhalation anthrax are consistent with this idea. They showed that animals treated for 5 or 10 days with penicillin remained free of disease while receiving the drug but died when the antibiotic was discontinued⁵.

Our experiments clearly demonstrate that more prolonged antibiotic treatment for a 30 day period results in statistically significant long-term survival after discontinuance of treatment. Seven of ten penicillin-, eight of nine ciprofloxacin-, and nine of ten doxycycline-treated animals survived. This result supports the hypothesis that treatment with antibiotics alone will be successful if the treatment continues until the level of retained persistent spores falls to less than the infectious dose. However, the five deaths in antibiotic-treated animals that were attributable to anthrax, particularly the animal in the doxycycline group that died 58 days after the initial aerosol exposure, directly support the concept that spores persist for prolonged periods in the host. The present data, taken together with the previously-reported treatment failure with a short course of antibiotics⁵, suggest that an even more prolonged course of antibiotics might have prevented all deaths from anthrax in these experiments. Although the duration of post-exposure antibiotic treatment required to prevent relapse is likely to be a function of the initial retained dose of spores, data are insufficient to determine accurately what the relationship is and how long treatment should be continued.

Our results did show that complete long-term survival, after discontinuance of antibiotics, occurred when post-exposure antibiotic-treatment was combined with vaccination. This result confirms the previous reports of Henderson et al.⁵ and Lincoln et al.¹⁴. The survival of all nine animals in the doxycycline plus vaccine group did not differ statistically from that of the animals treated with antibiotics alone. Nevertheless, antibiotics plus vaccination appears to be the optimal approach to post-exposure treatment of inhalation anthrax.

The serological response to the initial aerosol challenge was determined by measuring antibodies to the anthrax protective antigen. As shown in Table 2, none of the animals treated with penicillin, ciprofloxacin, or doxycycline alone developed evidence of an immune response to anthrax. This suggests that the antibiotic treatment inhibited significant bacterial growth and prevented the infection from developing. The only animals that seroconverted after the aerosol challenge were those that had been vaccinated. Seroconversion does occur in humans who recover from anthrax¹⁵⁻¹⁸. Vaccination likewise induces serological conversion in humans^{16,19} and non-human primates (unpublished data).

The development of an immune response was found to predict resistance to re-challenge. As seen in Table 3, the only animals resistant to a

second aerosol challenge with anthrax spores were those that had been vaccinated and had seroconverted. Antibiotics, which had protected animals against the initial infection, prevented the development of an effective immune response, leaving the animals susceptible to a subsequent infection. This agrees with a previous report where surviving animals treated post-exposure with antibiotics and by hyperimmunization with five doses of vaccine were protected upon re-challenge¹⁴. The protection afforded by vaccination before exposure is to be expected, based upon prior reports²⁰⁻²² as well as our experiments (unpublished data).

Thus, these results suggest that the optimal therapy for an unimmunized host exposed to an aerosol of anthrax spores consists of long-term suppressive antibiotic treatment together with vaccination to protect against both the initial challenge and any subsequent exposure.

References

1. F. M. LaForce, *Bull. N. Y. Acad. Med.* **54**, 956 (1978).
2. C. M. Dahlgren, L. M. Buchanan, H. M. Decker, S. W. Freed, C. R. Phillips, and P. S. Brachman, *Am. J. Hyg.* **72**, 24 (1960).
3. P. S. Brachman, *Annals. N. Y. Acad. Sci.* **353**, 83 (1980).
4. K. Vessal, J. Yeganehdoust, W. Dutz, and E. Kohout, *Clin. Radiol.* **26**, 471 (1975).
5. D. W. Henderson, S. Peacock, and F. C. Belton, *J. Hygiene* **54**, 28 (1956).
6. B. E. Ivins, S. L. Welkos, G. B. Knudson, and S. F. Little, *Infect. and Immun.* **58**, 303 (1990).
7. C. A. Gleiser, C. C. Berdjis, H. A. Hartman, and W. S. Gochenour, *Brit. J. Exp. Path.* **44**, 416 (1963).
8. R. E. Lincoln, D. R. Hodges, F. Klein, B. G. Mahlandt, W. I. Jones, Jr., B. W. Haines, M. A. Rhian, and J. S. Walker, *J. Infect. Dis.* **115**, 481 (1965).
9. W. S. Greenfield, *Eleventh Annual Report of the Medical Office of the Local Government Board, London* (1881-1882), pp. 207-238.
10. W.S. Albrink, S. M. Brooks, R. E. Biron, and M. Kopel, *Amer. J. Path.* **36**, 457 (1960).

FRIEDLANDER, WELKOS, PITT, EZZELL, WORSHAM, ROSE, IVINS, LOWE

11. W. S. Albrink, and R. I. Goodlow, *Amer. J. Path.* **35**, 1055 (1959).
12. J. M. Ross, *J. Path. Bact.* **73**, 485 (1957).
13. J. M. Barnes, *J. Path. Bact.* **59**, 113 (1947).
14. R. E. Lincoln, F. Klein, J. S. Walker, B. W. Haines, W. I. Jones, B. G. Mahlandt, and R. H. Friedman, *Antimicrobial Agents and Chemotherapy* - 1964. American Society for Microbiology, 759-763 (1965).
15. T. M. Buchanan, J. C. Feeley, P.S. Hayes, and P. S. Brachman, *J. Immunol.* **107**, 1631 (1971).
16. P. C. B. Turnbull, M. G. Broster, J. A. Carman, R. J. Manchee, and J. Melling, *Infect. and Immun.* **52**, 356 (1986).
17. T. Sirisanthana, K. E. Nelson, J. W. Ezzell, and T. G. Abshire, *Am. J. Trop. Med. Hyg.* **39**, 575 (1988).
18. H. Winter, and R. M. Pfisterer, *Schweiz. med. Wschr.* **121**, 832 (1991).
19. A. Johnson-Winegar, *J. Clin. Micro.* **20**, 357 (1984).
20. F. C. Belton, and R. E. Strange, *Brit. J. Exp. Path.* **35**, 144 (1954).
21. G. G. Wright, T. W. Green, and R. G. Kanode, *J. Immun.* **73**, 387 (1954).
22. H. M. Darlow, F. C. Belton, and D. W. Henderson, *Lancet* **II**, 476 (1956).

Battlefield Environmental Effects on Ultraviolet Atmospheric Sensing
Propagation

Dr. James B. Gillespie and Dr. David L. Rosen
U.S. Army Atmospheric Sciences Laboratory; SLCAS-BW-R
White Sands Missile Range, New Mexico 88002-5501

INTRODUCTION

Electro-optical systems design requires a knowledge of the effects of the atmosphere on the received signal. Although interest has been traditionally focused on infrared and longer wavelength systems, the availability of improved sources and detectors, the desire to utilize the increased resolution offered by shorter wavelength systems, and more secure limited distance propagation has led to a renewed interest in the shorter wavelength optical systems, extending into the solar blind region below 310 nm. Remote sensing of chemical and biological agents, constituent detection and identification and dosage quantification, and short range target detection are problem areas which may be applicable to UV technology. Our laboratory has instituted a basic research program to address the problems associated with ultra-violet propagation and detection in the atmosphere. This program consists of a theoretical UV propagation modeling effort and an experimental UV spectroscopy laboratory effort.

The UV propagation effort was originally setup to address the problem of atmospheric attenuation effects on the propagated beam and on the return signal. Closer investigation into the problem, however, revealed that background fluorescence by aerosols and gasses were a potential problem. Problems of fluorescence lifetimes, photon yields, atmospheric quenching, and particle size and shape effects were also encountered. The UV research work was organized into two distinct parts in order to address these problem areas: (1) a theoretical modeling research effort to explore spectral discrimination techniques, to model atmospheric effects on analysis of remote sensing systems, and to perform simulations of systems performance, and (2) experimental spectroscopy laboratory research to develop techniques to measure UV optical and fluorescence properties of possible target and background materials and to perform validation experiments of the theoretical models developed.

MODELING

A need has been demonstrated recently for an increased capability to model propagation and atmospheric backscatter at ultraviolet wavelengths. This arose from increased interest in the solar blind spectral region, UV lidar remote sensing, interpretation of fluorescent signatures, and UV transmission studies. Existing available models were primarily developed for the infrared and longer wavelengths. Their short wavelength capabilities were generally extrapolations into the UV region. For our requirements a UV propagation model must have the following features:

- (1) Model Propagation in the 185 to 700 nm spectral region for slant, horizontal, and vertical paths--primarily for the atmospheric boundary layer.
- (2) Include procedures which account for molecular scattering and absorption, aerosol attenuation, and sky brightness.
- (3) Include a lidar backscattering model which accounts for Rayleigh scattering, Mie scattering, and fluorescence.
- (4) The program itself must be an interactive, modular FORTRAN code capable of running on an IBM PC compatible computer and be menu driven and user-friendly.
- (5) The program must contain default values and have a standard atmosphere included.

Our propagation model, UVTRAN¹, differs from similar models in that it is designed to be used for UV and visible wavelengths only; so that band models and emission calculation need for IR wavelengths are not included. Molecular oxygen is also included and SO₂ and nitrogen oxides absorptions are also included. The aerosol model is parametrized in terms of visibility and wavelength², and the Molecular scattering is based on a formula by Penndorf³. These are given in equations 1 and 3, respectively.

$$\sigma_{aer}(\lambda) = \left\{ \frac{3.912}{V} - \sigma_{ms}(550) \right\} \left\{ \frac{\lambda}{550} \right\}^q \quad (1)$$

where λ is the wavelength, V is the visibility, σ is the extinction coefficient, and q is a parameter given by

$$q = 0.585 V^{\frac{1}{3}} \quad (2)$$

This aerosol model is a semi-empirical model based on short wavelength transmission data. The molecular scattering formula is given as

$$\sigma(\lambda) = c \lambda^{-4} \left\{ \frac{N}{N_0} \right\} (9.26 e18 - 1.07 e9 \lambda^2)^{-1} \quad (3)$$

where N is the number density of air molecules and N_0 is the number density under standard conditions.

A comparison of our model and the LOWTRAN 6 calculations for standard conditions is shown in Table 1. The calculations were made for wavelengths of 300, 550, and 700 nm with a visibility of 10 km. The rural model was used for the LOWTRAN 6 model. In this table t represents the transmittance.

TABLE 1. Transmittance Comparison						
Transmittance	300 nm		550 nm		700 nm	
	UVT	LOW	UVT	LOW	UVT	LOW
tms	.239	.233	.892	.890	.956	.957
tO3	.773	.768	.998	.998	.999	.999
taer	.013	.064	.206	.206	.347	.307
toth	.986	N/A	.997	N/A	N/A	.954
ttot	.002	.011	.182	.183	.331	.280

Baum and Dunkleman⁴ measured attenuation coefficients between 250 and 550 nm in Pasadena for a range of visibilities. They developed a strictly empirical model for UV propagation based on these measurements. Table 2 shows a comparison of aerosol attenuation coefficient, σ_{aer} , calculated by UVTRAN, LOWTRAN 6, Baum and Dunkleman's (BD) model. The difference between UVTRAN and the BD model suggest that the standard ozone concentrations during the Baum and Dunkleman measurements were higher.

TABLE 2. Comparison of σ_{aer} at 300 nm

Visual Range (km)	σ_{aer}		
	UVTRAN (km ⁻¹)	LOWTRAN 6 (km ⁻¹)	B D (km ⁻¹)
100	.137	.047	.213
40	.284	.150	.340
20	.476	.320	.553
10	.809	.662	.978
5	1.410	1.340	1.830

The major difference between UVTRAN and LOWTRAN 6 results is due to the different aerosol parametrization in the two models. The UVTRAN model is more directly based on empirical transmission data; it offers a simpler, yet more accurate parametrization of aerosol attenuation at visible and UV wavelengths than does the LOWTRAN parametrization.

The second part of our modeling effort involves examining the effect of atmospheric attenuation on the analysis of UV fluorescence signatures. Laser induced fluorescence (LIF) is used for remote sensing because the strength of the fluorescence enables chemical analysis to a kilometer or greater. Discrimination between chemical species is often difficult, however, because: (1) noise in the data, (2) background fluorescence of aerosols and gasses, (3) quenching, and (4) fluorescence spectra are in general broad and somewhat featureless. There are two primary analysis techniques for fluorescence spectra: Least squares (LS) fitting and rank annihilation - factor analysis (FARA). LS requires knowing specifically every substance that is present and is sensitive to noise. This is generally not practical for situations outside the laboratory. FARA is an eigenanalysis algorithm for discriminating between species in a mixture using the excitation-emission matrices (EEM) of LIF spectra. This technique is more applicable to lidar returns because it is insensitive to unknown compounds. This technique does require, however, that the number of excitation wavelengths exceed the number of fluorescing compounds. A tunable laser is therefore necessary for this type of work. We performed a computer simulation to investigate the effect of atmospheric attenuation on FARA⁵. We used UVTRAN to generate the atmospheric spectral attenuation profile. We used the spectra of three amino acids, tryptophan (TRY), phenylalanine (PHE), and tyrosine (TYR). We applied FARA to a simulated mixture of TRY and PHE. The effective concentrations for TRY and PHE were 20 and 1 respectively. Excitation wavelengths from 260 to 290 nm and emission wavelengths from 300 to 440 nm were used in our

calculations. Define $T(\lambda;R)$ as the transmission of the atmosphere at wavelength λ over a distance R , $G(R)$ as a wavelength dependent term which normalizes the geometry of the lidar apparatus, N' as the EEM of the mixture after inclusion of atmospheric effects. We then have:

$$N_{ij} = T(\lambda_i;R) T(\lambda_j;R) G(R) \dot{N}_{ij} \quad (4)$$

where the subscript i indexes the emission wavelengths, and the subscript j indexes the excitation wavelengths. Define D as the reference EEM for the target compound, and c as a hypothetical effective concentration. The target compound in our simulation was TRY. An error matrix, M , is defined in FARA as⁶:

$$M(c) = N - c D \quad (5)$$

If M^t is the transpose of M and r is the number of fluorescing compounds in the mixture then the r th largest eigenvalue, $C(c)$, of $M(c)^t M(c)$ reaches a minimum at some value of c which we define as the apparent concentration, c_a . In a laboratory where there is no atmospheric attenuation, the apparent concentration is the same as the effective concentration c_r , and $C(c_r)$ is zero. Calculations were done for remote measurements with extinction that corresponds to visibilities of 10, 23, and 50 km at sea level.

The eigenvalue $C(c)$ is plotted in figure 1 for four distances at a visibility of 23 km. The changes in spectral shape affected these eigenvalue curves in two obvious ways. First, the relative depth of the eigenvalue curve at the minimum has decreased. Second, c_a has become smaller than c_r , so that $c_a = c_r = 20$ only in the laboratory. A plot of the reduction factor for apparent concentration, c_a/c_r , is shown in figure 2. These figures show that atmospheric attenuation becomes significant even at small distances (<0.2 km). This type of analysis has shown that atmospheric degradation of fluorescent spectral signatures obtained by UV lidar can strongly impact on the FARA method. The spectrum must be deconvolved from the atmospheric extinction profile in order to properly interpret spectral returns. We have recently developed qualitative detection algorithms which show that for identification only, this is not as serious a problem as it is with the quantitative FARA analysis.

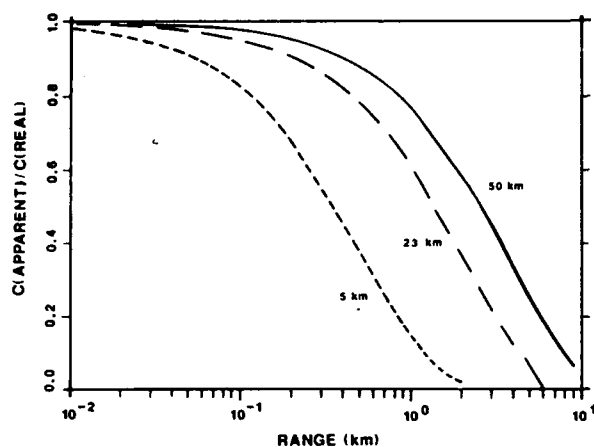


Figure 1. The ratio of the apparent concentration to the real concentration versus range for 4 visibilities.

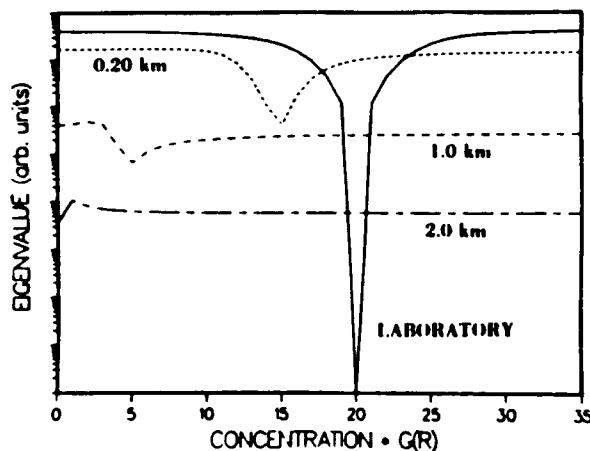


Figure 2. Eigenvalue versus range for 4 visibilities.

Using the methodology of the FARA method above, we performed a modeling simulation to determine the effects that shot noise and sky radiance would have on a typical UV fluorescence lidar system. We chose a laser pulse energy of 0.1J, a receiver mirror diameter of 0.6m, a receiver field of view of 2Mrad, a spectral bandwidth of 2.5nm, a transmitter efficiency of 0.55, a receiver efficiency of 0.54, and averaged over 6000 pulses. We applied this to a problem of detecting NO_2 in an atmospheric background that contains fly ash as a background fluorescing material. We used a discrimination method known as the overlap index⁷ to identify the NO_2 in the EEM of the lidar return. The overlap index is given as

$$p = \sum_{i=1}^I |u_i \cdot x|^2 \sum_{j=1}^I |v_j \cdot y|^2 \quad (6)$$

The technique requires several excitation wavelengths and the lidar return emission spectrum for each of these wavelengths. Figure 3 shows the emission spectra for NO_2 (dashed lines) and for the fly ash (dotted line) superimposed on a hypothetical lidar return (solid line) for a 2:1 mixture of NO_2 to fly ash for an excitation wavelength of 457.9 nm. The signal to noise ratio for (A) is 16 while the signal to noise for (B) is 3. Figure 4 shows the overlap index, p , for this example. Figure 5 shows a three dimensional projection of the EEM for the for the lidar return at 0.124 km: (A) with only intrinsic photon noise and no sky radiance and (B) with both intrinsic photon noise and nighttime sky radiance.

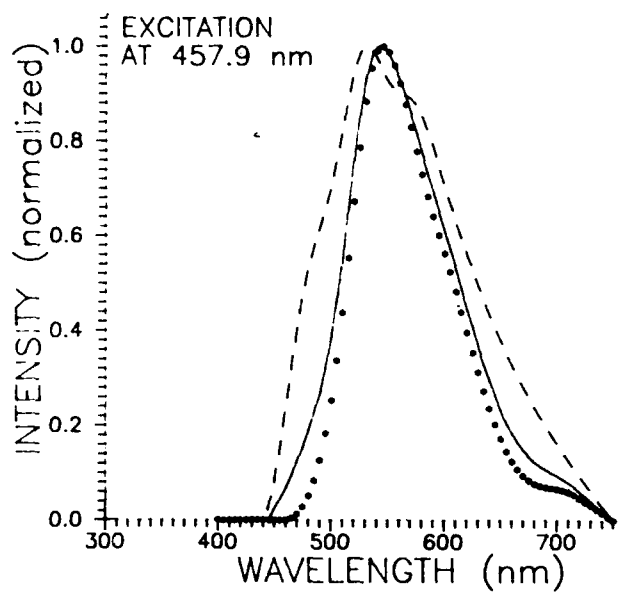


Figure 3. Emission Spectra for NO, fly ash, and the mixture.

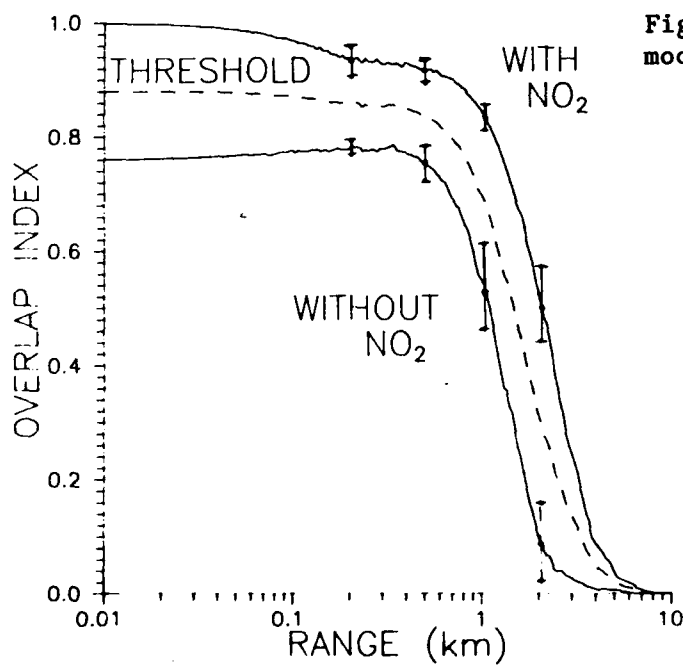


Figure 4. Overlap index for the moonless night case.

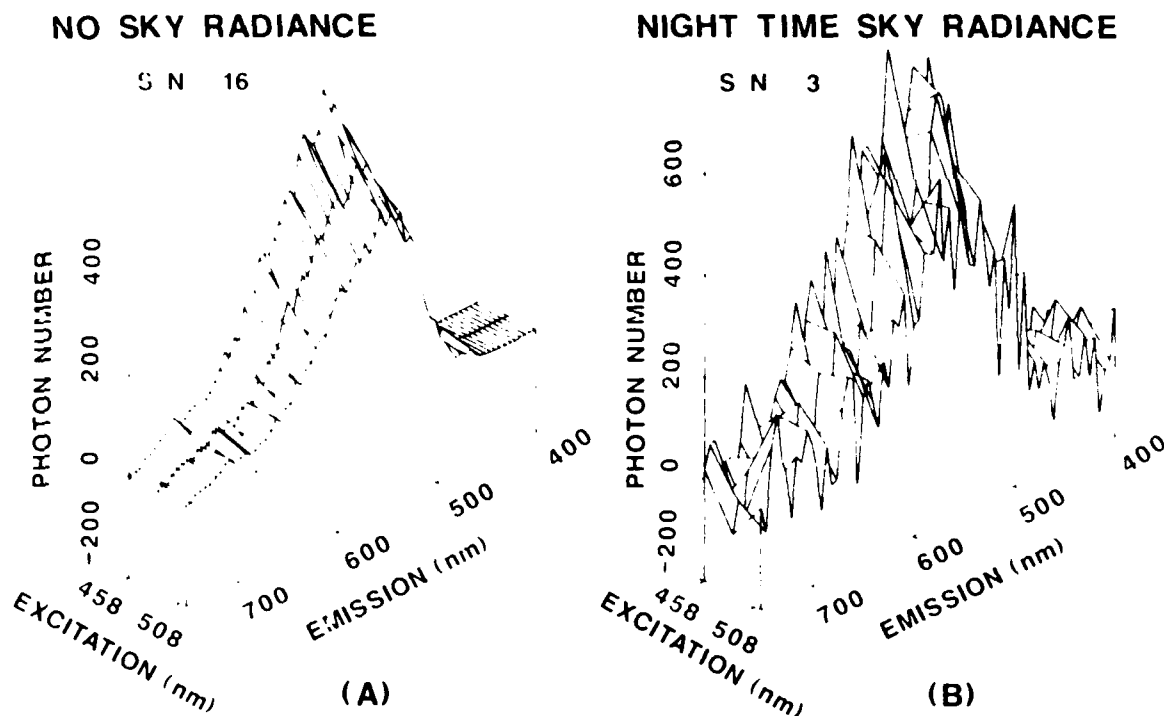


Figure 5. Simulated lidar return for (A) no sky radiance (S/N=16) and (B) for a nighttime sky radiance (S/N=3)

For a moonless night a range of one kilometer should be obtainable. For a moonlit night, a range of 200m is easily obtainable. For a daytime situation we found that fluorescence lidar will not work for the situation that we modeled.

LABORATORY SPECTROSCOPY

Because of the lack of information available on fluorescence properties of aerosols, a UV spectroscopy laboratory was instituted to characterize aerosols and gasses and to verify our modeling. The following laboratory research capabilities are available in our laboratory:

1. Singly levitated particle spectroscopy
2. Aerosol Chamber and gas cell Spectroscopy
3. Diffuse Reflectance Spectroscopy of powdered materials
4. EEM fluorescence spectra measurements
5. Two-photon excitation, degenerate and non-degenerate
6. Time resolved fluorescence measurements

GILLESPIE and ROSEN

The singly levitated particle spectroscopy has the capability to characterize the following properties of particles:

1. Obtain fluorescence spectra of 0.5 to 80 μm particles
2. Measure fluorescence lifetimes and photon yield
3. Determine particle size, weight, and density
4. Measure refractive index of spherical particles
5. Determine thermal and physical properties
6. Measure humidity and particle growth effects

Recent experiments were performed on singly levitated micrometer size droplets to determine the effect of absorption on the morphology dependent resonances of spherical particles⁸. A glycerin droplet was charged and injected into a hyperbolic electrodynamic quadrupole trap⁸. A particle of pure glycerin was illuminated with an argon ion laser while it slowly evaporated over a period of several hours. The scattering was observed at 90 degrees and recorded on a strip chart recorder. Subsequent particles of glycerin were doped with a strong absorber to change the imaginary part of the refractive index. The scattering spectra of these particles was recorded. The quality factor, Q , of selected resonances was experimentally determined. Figure 6 shows the experimental scattering data for three particles of different imaginary refractive index. The experimental data is superimposed on the theoretical Mie curve for $n = 1.4722$ and $k = 0.0, 1.1\text{e-}6$, and $2.8\text{e-}6$. Figure 7 shows the absorption effect on the intensity of the TE_{20}^{14} resonance.

Laser spectroscopy measurements have been performed on samples of interest to the Drug Enforcement Administration (DEA). These characterizations consist of fluorescence emission spectra of illicit drugs excited at 355 nm and 256 nm, fluorescence photon yields and lifetimes for 308 nm and 248 nm laser light for selected solvents used in illicit drug manufacture, and multi-photon induced fluorescence spectra of solvent vapors. These data are available to qualified users upon request.

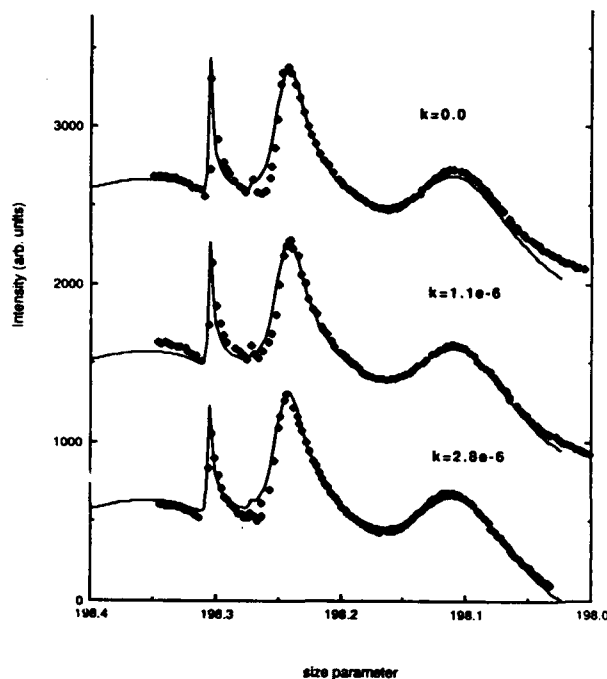


Figure 6. Experimental scattering data for $n=1.4722$ and $k=0.0$, $1.1e-6$, and $2.8e-6$.

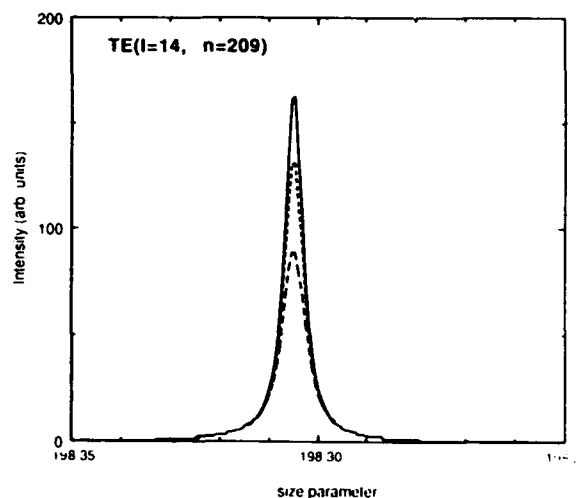


Figure 7. The effect of absorption on the TE(1-14, $n=209$) resonance.

The diffuse reflectance spectroscopy facility has the capability to obtain absorption spectra of powdered materials and to determine the imaginary part of the refractive index of substances which exist only in powdered form. This technique has been pushed down to $0.22 \mu\text{m}$ in wavelength by our laboratory. Data on carbon, kaolin clay, iron oxide, and many different minerals is available from measurements we have made with our diffuse reflectance setup⁹⁻¹¹.

CONCLUDING REMARKS

Modeling of atmospheric effects on analysis of interpretation of spectral returns from a hypothetical UV lidar system has demonstrated the applicability and the limitations of UV lidar for remote chemical detection. The modeling indicates that such a system is feasible for nighttime operation, and by applying analysis tools such as Factor - Analysis - Rank - Annihilation, identification of target fluorescing agents can be made to ranges of 1km or more. However, a tunable laser is required for such operation.

GILLESPIE and ROSEN

Experimental laboratory measurements have been made on singly levitated microparticles and on particulate and gaseous samples of interest to the counter narcotics effort. The results of the counternarcotics measurements can be made available to qualified organizations upon request.

REFERENCES

1. E.M. Patterson, and J.B. Gillespie, "Simplified Ultraviolet and Visible Wavelength atmospheric Propagation Model," *Appl. Opt.* 28, pp425-429 (1989).
2. P.W. Kruse, L.D. McGlauchin, and R.B. McQuistan, *Elements of Infrared Technology*, Wiley, New York (1963).
3. R. Penndorf, "Tables of the Refractive Index for Standard Air and the Rayleigh Scattering Coefficient for the Spectral Region between 0.2 and 20 μm and Their Application to Atmospheric Optics," *J.Opt. Soc. Am.* 47, 176 (1957).
4. W.A. Baum and L. Dunklema, "Horizontal Attenuation of Ultraviolet Light by the Lower Atmosphere," *J.Opt. Soc. Am.* 45, 166 (1955).
5. D.L. Rosen and J.B. Gillespie, "Atmospheric Extinction Effect on Analysis of UV Fluorescence Signatures," *Appl. Opt.* 28,4260-4261,1 Oct 1989.
6. E.R. Malinowski and D.G. Howery, *Factor Analysis in Chemistry*, Wiley, New York (1980).
7. D.L. Rosen and J.B. Gillespie, "Fluorescence Lidar Detection with Shot Noise and Sky Radiance," *Appl. Opt.* in press (1992).
8. M. Essien, J.B. Gillespie, and R.L. Armstrong, "Observation of Suppression of Morphology-dependent resonances in singly levitated micrometer sized droplets," *Appl. Opt.* in press (1992)
9. J.B. Gillespie and J.D. Lindberg, "Measuring Absolute Diffuse Reflectance in the Ultraviolet," *Appl. Opt.* in press (1992).
10. J.B. Gillespie and J.D. Lindberg, "Ultraviolet and Visible Imaginary Refractive Index of Strongly Absorbing Atmospheric Particulate Matter," *Appl. Opt.* in press (1992).
11. J.B. Gillespie and J.D. Lindberg, "Seasonal and Geographic Variations in Imaginary Refractive Index of Atmospheric Particulate Matter," *Appl. Opt.* in press (1992).

Landmine Detection by Backscatter Radiation Radiography

Ms. Janine Guadagno,* Dr. Ramachandra K. Bhat, Sgt. John Burchanowski
U.S. Army
Belvoir Research, Development and Engineering Center
Fort Belvoir, Virginia 22060-5606

A significant factor for battlefield survivability is the ability of the ground fighting forces to traverse the battlefield. The inability to detect nonmetallic mines has been a major obstacle in this arena (Figure 1). To address this deficiency, the Belvoir Research, Development and Engineering Center (BRDEC) assembled an expert panel in 1985¹ to investigate the use of ionizing radiation techniques for detecting landmines and to identify technological advancements that would alter the negative assessment of this technique made in the U.S. Army Mobility Equipment Research and Development Center Report 2134.² The primary focus of the 1985 workshop was on the detection of antivehicular (AV) nonmetallic landmines. To evaluate the techniques, the following set of operational requirements for AV mine detection units was developed:

Operational Requirements

<u>Characteristics</u>	<u>Desired Capability</u>	<u>Useful Capability</u>
Forward speed	1.34 meters/second	0.23 meters/second
Path width	3.05 meters	1.83 meters
Detection probability	0.99	0.90
False alarm probability	0.00075	0.00328
Power requirement	< 5 kilowatts	< 100 kilowatts
Maximum depth of burial for detection	20.32 centimeters	5.08 centimeters

Based on these requirements, and due to technological advancements made in high-intensity linear scanning x-ray sources and portable computer systems capable of high speed data and image processing, the panel considered the x-ray backscatter imaging approach to have a substantial possibility for application to the nonmetallic mine detection problem and to be the most promising of the techniques explored during the workshop.¹ Upon award of an Army contract, research on mine detection by the x-ray backscatter technique was initiated at the University of Florida in 1986.³ In 1988, BRDEC's Radiation Research Team (RRT) began an in-house research effort to augment the findings made by the University of Florida and determine the practical requirements for an AV nonmetallic mine detection unit utilizing x-ray backscatter radiography.

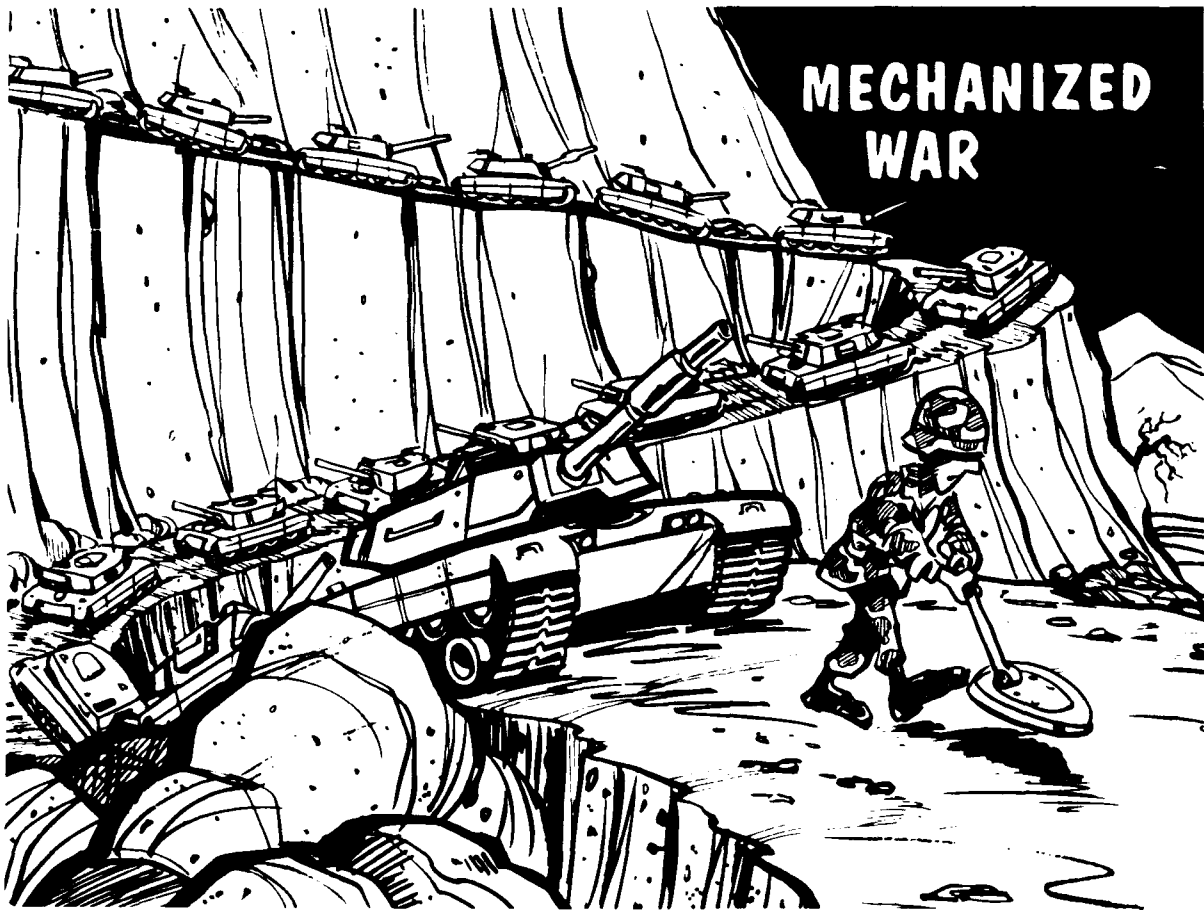


Figure 1. Mine Detection Capabilities

PRINCIPLES OF METHODOLOGY

Landmine detection through x-ray backscatter radiography is dependent upon the interaction of the x-ray photon with the material upon which it impinges. There are three main types of interactions that may take place: photoelectric absorption, Compton scattering, and pair production. The relative probabilities of these interactions occurring depends on the energy of the photon and the atomic number (Z) of the incident material. For low to intermediate photon energies, less than 300 keV, the photoelectric interaction coefficient in the explosives of mines (average Z of explosive material is 7) is significantly less than it is in typical soils (average Z is about 11 or 12). This difference in photoelectric absorption between mine and soil increases as the x-ray photon energy decreases. The primary scattering mechanism is the Compton scatter effect. The Compton scatter coefficient is nearly the same for the two materials. Photons which are not absorbed are scattered, and some of the scatter is backscatter which will travel up through the soil to the detector. Therefore, as mine replaces soil, the number of photons that are absorbed is reduced, and consequently the number of photons that are backscattered and escape from the mine and soil is increased. Thus, the photoelectric effect provides contrast between the mine and soil materials while the Compton effect carries the signal to the detector. The net result is an enhanced backscattered signal when the x-ray beam is incident on the mine.

A collimated x-ray beam, perpendicularly incident on the soil surface, is rastered across the soil surface to provide an image consisting of pixels that are characteristic of the photon interaction characteristics of the material through which the beam and its scattered components pass. Collimation of the x-ray source is required to allow for smaller pixel definition which is critical to the resolution of the mine image.

The optimal x-ray beam energy is determined by the physics of the Compton interaction. Low energy photons are unable to penetrate the upper soil layer because of the dominance of the photoelectric effect. High energy photons are able to penetrate more deeply, but because of their higher fractional energy loss and the forward peaking of the Compton scattering distribution as energy increases, they are not effective in the mine detection technique. Intermediate x-ray energies in the range of 100 to 200 kVp are used because, in this range, the optimum compromise between the required soil penetration, necessary to image a mine buried to a depth of

3 inches, and enhancement of the mine soil photon backscatter differences is achieved.

By collimating the detector array, primary scatter photons—photons which are scattered once directly back to the detector array—are eliminated. Primary scatter photons reflect very little of the mine soil difference and are predominantly from the soil above the buried mine. Primary scatter photons represent the larger fraction of the backscattered photons and will serve to mask the detector response to the buried mine. Secondary photons—photons which undergo multiple scattering in soil or mine before reaching the detector array—amplify the difference between the materials because they must survive multiple opportunities of photoelectric absorption (Figure 2). It is these secondary scatter photons that are used for mine imaging.

Monte Carlo codes were developed at the University of Florida and the United States Military Academy to examine the dependence of the response of the detection system to incident beam energy, beam angle of incidence, detector geometry, soil type and density, mine geometry, and depth of burial of the mine. These codes serve as the theoretical basis for the RRT research efforts.

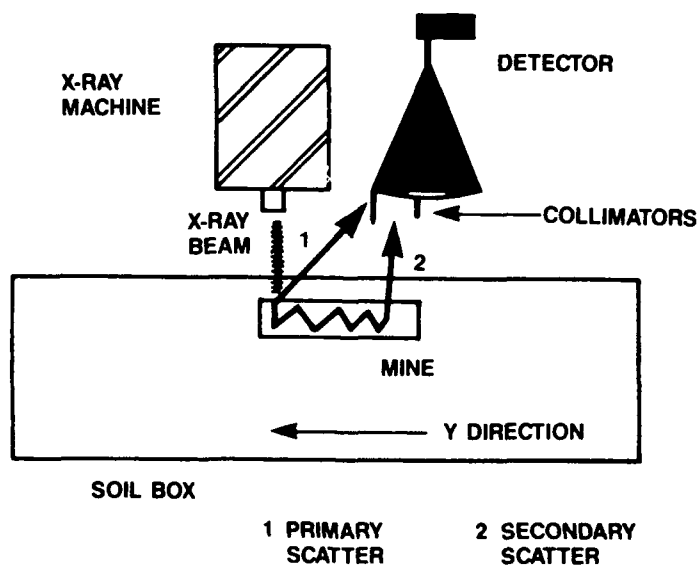


Figure 2. Principles of Scatter Radiation Radiography

EQUIPMENT

The RRT's x-ray backscatter research system consists of an OX-250 Industrial Radiographic unit, a 5 foot by 5 foot by 2 foot soil box, a Unidex 11 soil box positioner, a personal computer with an IEEE board, an Ortec high voltage power supply, an Ortec quad counter, and four 3 inch by 1 inch sodium iodide (NaI) detectors (Figure 3). The x-ray head remains stationary during the experiments and ground movement is simulated by the rastering of the soil box under the x-ray beam. The RRT developed a set of radiation protection safety protocols for the long continuous x-ray unit run times required in the experimental configuration. These protocols include the use of personal dosimeters, continuous monitoring with Geiger counters, and remote viewing devices.

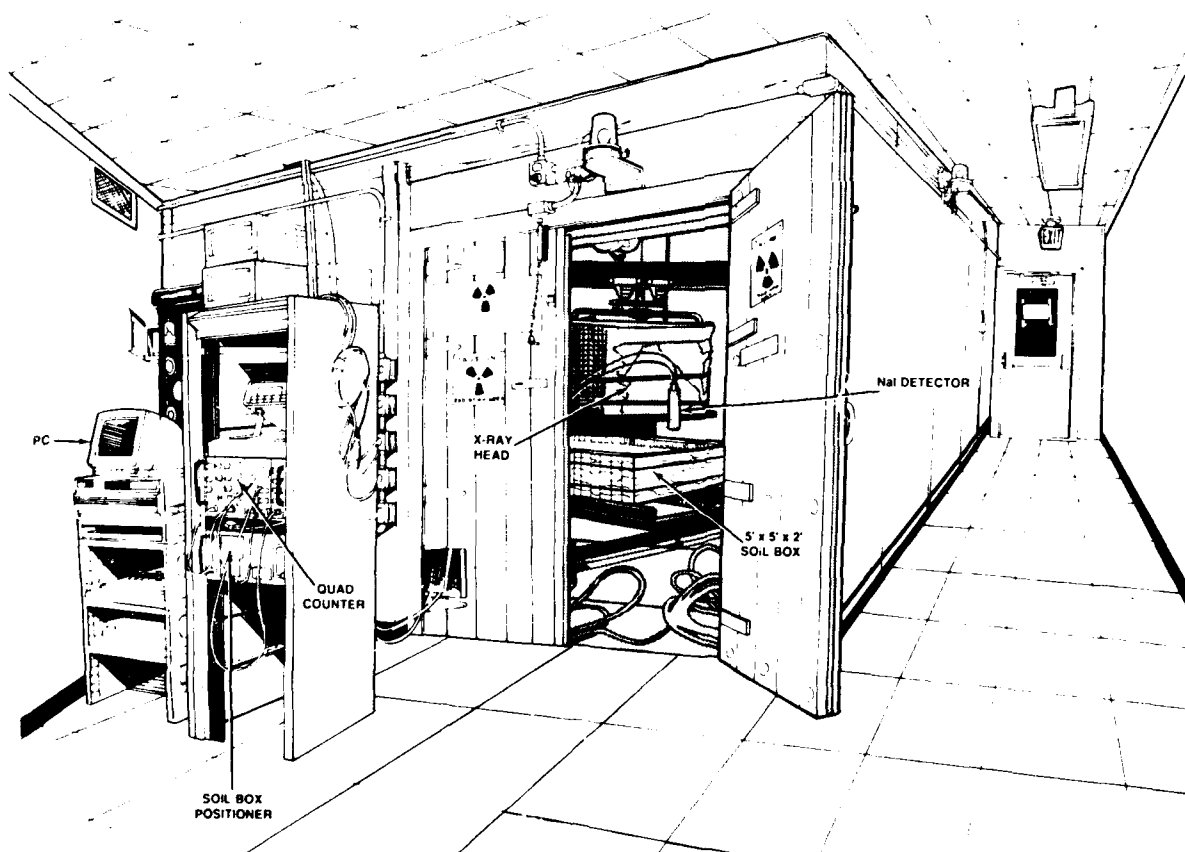


Figure 3. X-ray Backscatter Research Facility

CALIBRATION

The first practical consideration for the detection of mines using x-ray backscatter radiography addressed by the RRT is the calibration procedures required for the initial set-up and daily maintenance. The fielded backscatter unit is projected to have two large area panel detectors placed in front of and behind the rastering x-ray beam (Figure 4). Multiple detector imaging will require an averaging of the detector responses; therefore, the responses of the detectors must be balanced. Calibration procedures are required which will insure the multiple detectors are balanced and are not operating beyond their saturation level. Our research using NaI detectors shows that without calibration, detectors with identical operational parameters and configurations do not have the same response rate under similar exposure scenarios. Before calibrating, images produced by two detectors under similar exposure scenarios show poor balancing of responses (Figure 5). After calibration, the images produced by the two detectors under similar exposure scenarios demonstrate the balance of response required for multiple detector imaging (Figure 6). The typical energy range of the backscatter photon reaching the detectors is between 0 and 120 keV. Detector response at less than 20 keV includes electronic noise and reduces image clarity. For calibration of the detectors, an Americium-241 (Am-241) source is taped to the center of each detector. Using an oscilloscope and a single channel analyzer (SCA), the coarse gain is adjusted for each detector to insure the response is not over or under compensated. After the coarse gain is set, the lower level discriminator (LLD) is adjusted, using an oscilloscope, to eliminate the detection of energies in the noise range of 0 to 20 keV. Finally, the detector responses are matched by adjusting the fine gain on the SCA until the response rate, as recorded on the quad counter, is similar for all the detectors. The average response rate for the detectors is recorded and used in the daily calibration tests. These initial calibration procedures will have to be conducted for all multiple detector units and should be performed by the system operator so that the individual understands the limitations of their counting system. Daily calibration procedures for the detector train, to be included in the operation of the fielded unit, will involve the placement of a source on each detector face to check the response rate. If the response rate is found to deviate more than 1σ from the mean rate of the detectors' response, as determined during the initial calibration procedures, adjustment to the SCA fine gain will be required to correct for the drift.

Our research also shows the importance of checking the diameter of the incident beam and the location of the beam. Too large of beam size makes detector collimation extremely difficult and reduces the resolution in the image. Our research indicates that measurements of the x-ray beam strength, size and angle of incidence are critical to imaging and will require initial and periodic monitoring of the fielded unit to insure the reliability of the data. In our experiments, beam size and angle of incidence at the soil surface were determined by film exposures.

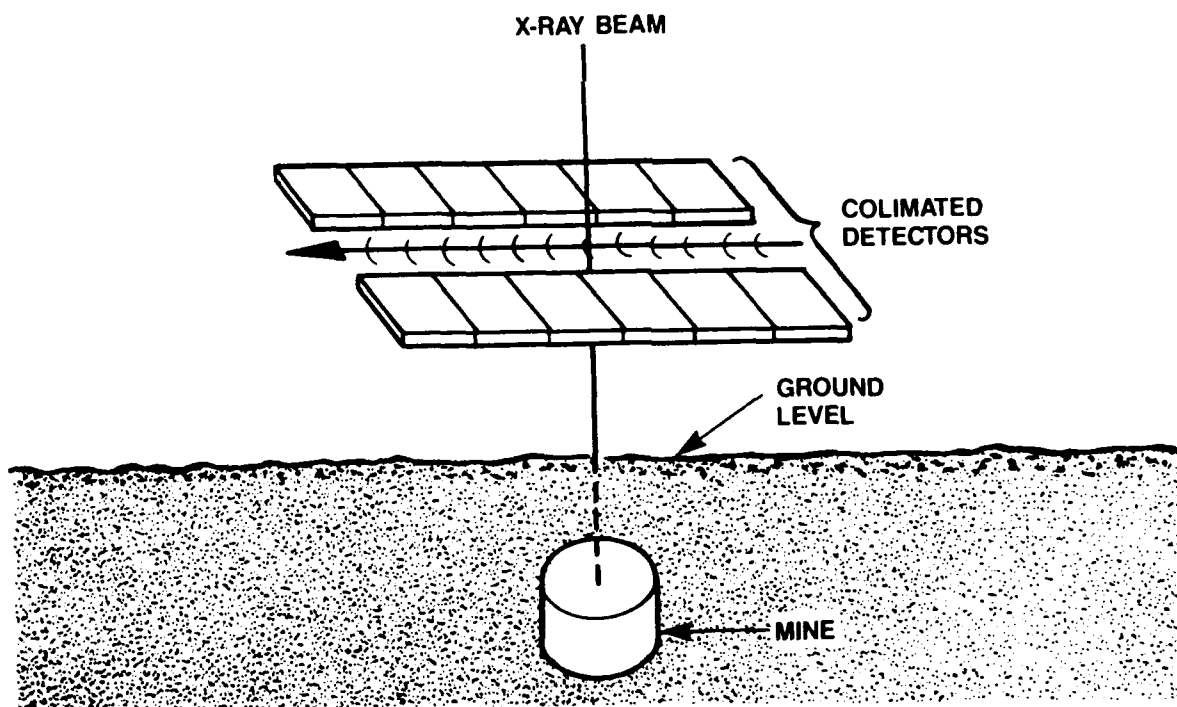
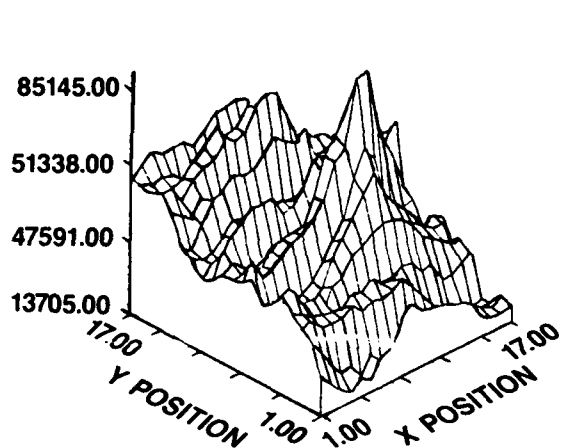
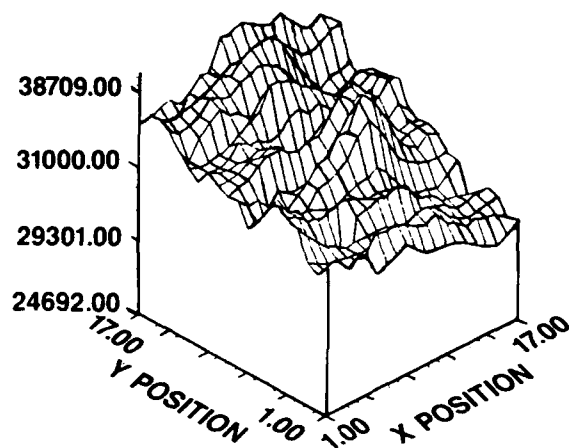


Figure 4. End Item Backscatter Field Unit Detector/X-ray Configuration

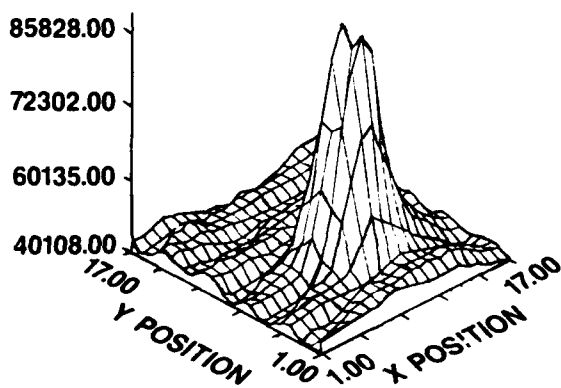


**Detector 3 In Position B
Unbalanced**

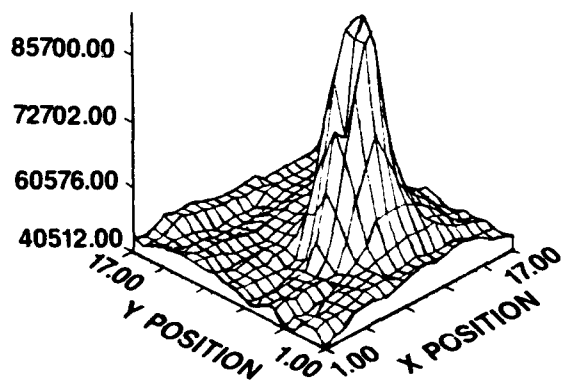


**Detector 4 In Position B
Unbalanced**

Figure 5. Before Calibration



**Detector 3 In Position A
Balanced**



**Detector 4 In Position A
Balanced**

Figure 6. After Calibration

RESULTS

Monte Carlo calculations predict that without detector collimation, mines can be detected only to depths of 2.54 centimeters; and with detector collimation, mines can be detected to a depth of 7.62 centimeters. Figure 7 is the Monte Carlo generated prediction of a buried mine image without detector collimation.³ Our experiments verify the Monte Carlo prediction that mines buried to depths of 5 centimeters cannot be detected without detector collimation; however, with detector collimation, a mine image is clearly produced at this depth (Figure 8). With optimal detector collimation we are able to image mines buried to a depth of 9 centimeters (Figure 9). Our research also demonstrates the practical consideration of height sensitivity to imaging using collimated detectors. Detector collimation length is strictly a geometry-driven parameter which is based upon the height of the detector from the soil and the distance of the detector from the x-ray source. The technique of similar triangles is employed to determine the optimal collimator length. Figure 10 demonstrates this technique. A maximum scattering depth of 10.16 centimeters is selected for the calculations to exclude single scattered photons from reaching the detector. To determine the sensitivity of the imaging process to fixed height variations, images were made of a mine buried 2 inches below the soil using the fixed length collimator at detectors heights varying from 48 to 79 centimeters (Figure 11).

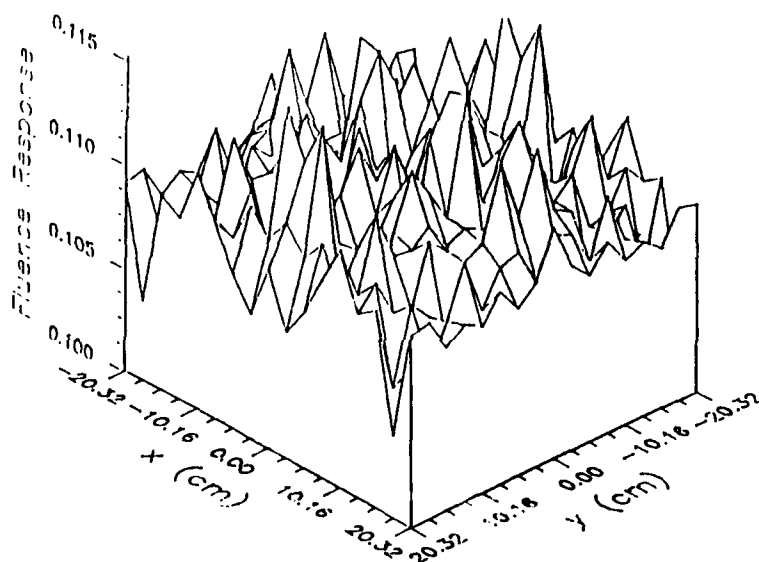
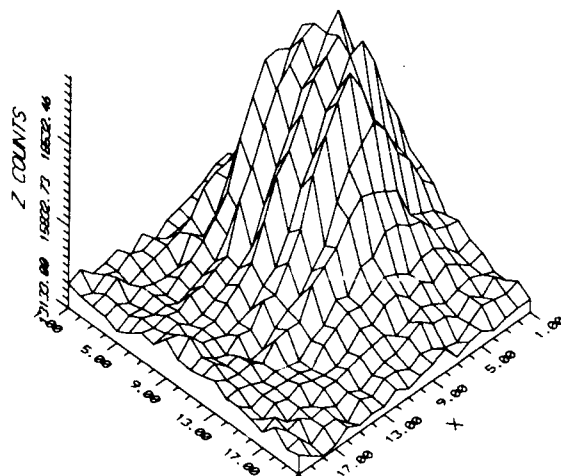
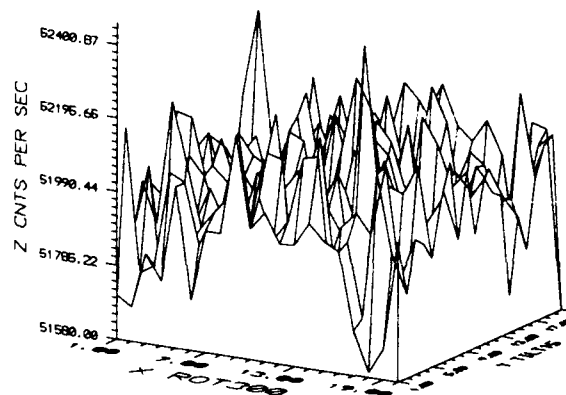


Figure 7. 5 cm Buried Mine Uncollimated Detector Monte Carlo Generated

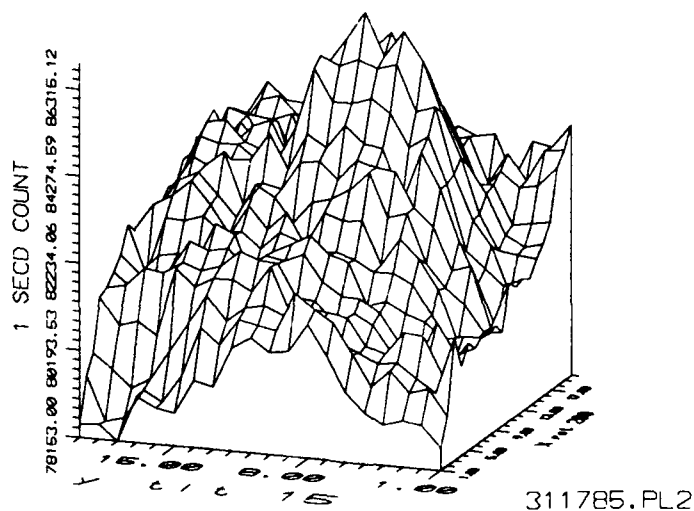


**5 cm Depth Mine;
Collimated Detector**



**5 cm Depth Large Mine;
Uncollimated Detector**

Figure 8. Images Produced Using Collimated and Uncollimated Detectors



INEL 12 Mine, 9 cm Depth of Burial

Figure 9. Image at Maximum Depth of Burial

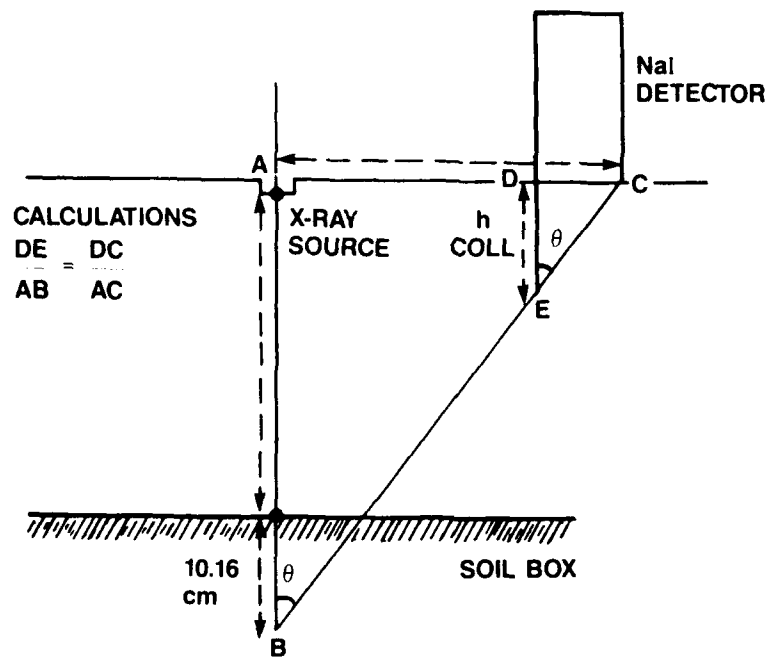
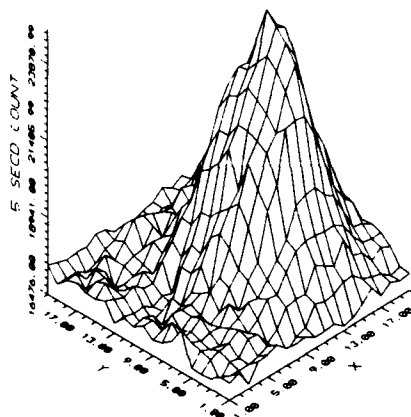
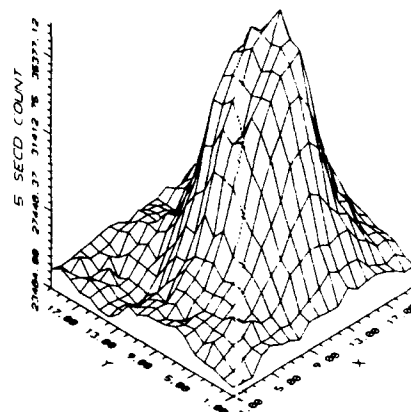


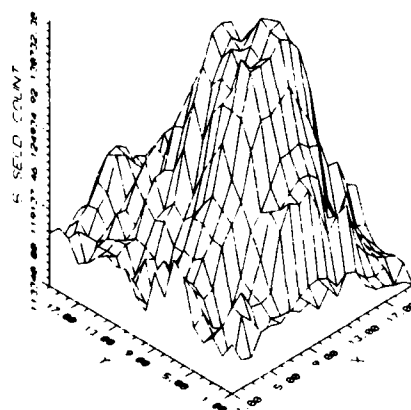
Figure 10. Similar Triangle Technique to Determine Detector Collimation Length



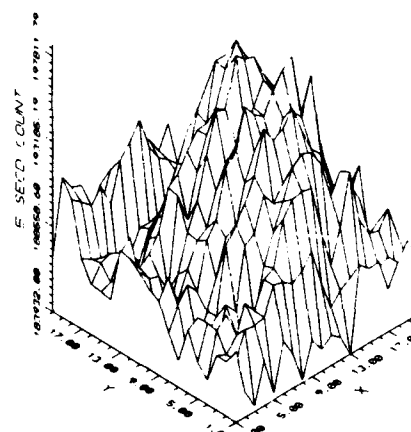
**Large Mine 5 cm Buried
48 cm High H455/271.1**



**Large Mine 5 cm Buried
58 cm High H455/274.1**



**Large Mine 5 cm Buried
68 cm High H455/279.1**



**Large Mine 5 cm Buried
79 cm High H455/27C.1**

Figure 11. Height Variations Using Fixed Length Collimeter

CONCLUSIONS/FUTURE WORK

Our work with multiple detector arrays identifies critical calibration procedures necessary for nonmetallic mine imaging using the x-ray backscatter technique. These procedures will need to be included in the operation manual for the field unit. Our research indicates that mine images obtained using a fixed length collimator are not very sensitive to height variations above the soil, providing the height above the soil is kept constant throughout the imaging process. There is also a range of heights above the soil to which the detectors can be set without significant image degradation. Additional study is required to determine the effect of height variations which occur during the image process. These variations could be caused by surface artifacts or surface irregularities and may badly distort the mine image making it unrecognizable. Research on image enhancement and acquisition is required and may be able to compensate for some loss of imaging sensitivity due to height variation during the imaging process.

REFERENCES

1. Workshop Report: Robert B. Moler, *Nuclear Techniques for Mine Detection Research*, July 22-25, 1985, Lake Luzera, NY (Belvoir Research, Development and Engineering Center, Fort Belvoir, VA).
2. Report 2134: Fredrick L. Roder, CPT Richard A. von Konynenburg, *Theory and Application of X-ray and Gamma-Ray Backscatter to Landmine Detection*, March 1975 (US Army Mobility Equipment Research and Development Center, Fort Belvoir, VA).
3. Technical Progress Report: Dr. Jacobs, *Landmine Detection by Scatter Radiation Radiography*, May 7, 1986 (University of Florida, Gainesville, FL).

Optical Neural Network Device Architecture

Dr. N. Gupta,* Mr. J. Pham, Dr. G. J. Simonis, Dr. R. P. Leavitt, and Mr. M. R. Stead
U.S. Army LABCOM, Harry Diamond Laboratories
2800 Powder Mill Road
Adelphi, MD 20783-1197

1. Introduction

The problem of pattern recognition is of utmost importance to the U.S. Army for aided target recognition (ATR). Besides military applications pattern recognition has many civilian applications, such as speech recognition, machine recognition of handwriting, robotic vision, communication systems, automated control, design and manufacturing, radar data analysis, etc.

So far the human brain is the most outstanding system for performing pattern recognition. It can recognize a pattern even with partial or incomplete information and in relatively ill-defined situations. On the other hand, a digital computer, which is very fast in number manipulation and other serial operations, is very inefficient in pattern-recognition tasks. It would require an enormous memory, a very large knowledge base, and a rule-based expert system to do relatively simple ATR tasks using traditional methods on electronic computers. The emerging field of artificial neural networks (NN) technology is based on the present knowledge of how a biological brain functions as a parallel processor using a large number of relatively simple but slow (millisecond) processing elements (neurons), working in concert with each other with a massive number of interconnections.¹⁻³ Such an NN can be used to perform processing at several levels in ATR problems—from input data preprocessing (tasks such as data recovery, noise removal, signal-to-noise-ratio (SNR) enhancement, signal and/or image segmentation, etc) through final target identification. Much work is going on in the software and hardware implementations of NN's. Both electronic and optical approaches are being used for hardware implementations.

Most of the optical implementations have been based on use of a vector-matrix multiplication concept to simulate neural functioning.⁴⁻⁶ Most of these use a free-space architecture and require many optical components which cannot be integrated

monolithically. We have implemented an optical neural architecture in GaAs which is very compact and monolithic and is inspired by the natural fiber-like structures common in all biological neurons. It is based upon a newly emerging basic technology of semiconductor quantum wells and superlattices, which are being investigated for developing novel optical modulators and switches.⁷⁻⁸ The device emulates the basic neural decision-making process. It accepts multiple inputs, applies different adjustable weights to them, sums these weighted inputs, and uses an optoelectronic thresholding switch to make a binary decision. So far the architecture has been successfully implemented to design and test a two-input/one-output (2-to-1) neuron and is being expanded to include more complex single neurons and an NN. During the course of this implementation a very-high-dynamic-range (250 dB/cm) optoelectronic modulator and a highly nonlinear optoelectronic (25 dB) thresholding device have been designed and tested. A very high on/off ratio (2300) which is mostly input limited is obtained for the neuron device. The research described in this paper was carried out starting from the basic physics research to a device design application in our laboratories under an in-house research program.

2. Requirements of Neural Networks

The success of biological systems appears to be due to their complex and massively parallel architectures. A 3-lb human brain consists of over 10 billion neuron cells communicating among themselves through networks of over 100 trillion synaptic interconnections. The working principles of neural functions are essentially the same from a squid's brain to a human brain. The main difference is in the number of cells and interconnections.

Each neuron has a relatively simple structure and functioning method. A typical neuron consists of a cell body, ranging from about 5 to 100 μm in diameter, from which emanates one major fiber, the axon, and a number of fibrous branches, the dendrites. The axon carries the outgoing signal from the neuron and near its end it usually branches out extensively. The dendrites receive the input signals from other neurons, where they form a contact, called a synapse. A weight is applied to each signal, and the cell body sums these weighted signals and performs a simple thresholding operation based on this sum to send a signal on its axon or not. In general, one neuron receives input from thousands of other neurons through its dendrites and similarly feeds its output to thousands of other neurons through its branching axon. The time scale involved is relatively slow (\sim milliseconds). Figure 1(a) shows typical biological neurons and figure 1(b) shows schematics of a neuron's operational principle. The basic neural operation may be summarized as follows:

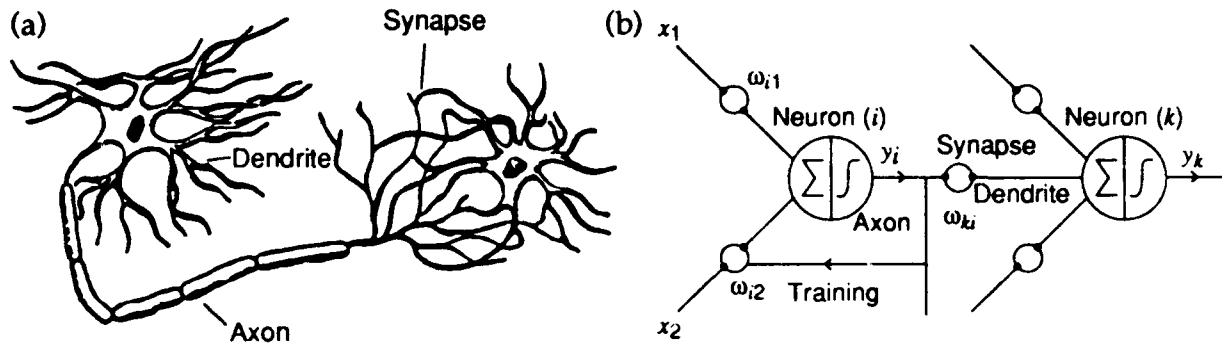


Figure 1. Neuron systems: (a) two biological neurons, and (b) schematic of neural operational function showing inputs (x_j), weights (w_{ij}), summing, thresholding, and output (y_i) for neuror (i).

- A single neuron has multiple inputs, multiple interconnections, and a single output.
- Each input has an adjustable synaptic weight associated with it which may have either positive or negative value.
- The neuron must sum weighted inputs and exhibit a nonlinear transfer function in issuing an output signal.

An artificial NN consists of a massively interconnected network of such processing elements. For mathematical simplification, the network is divided into a number of layers; each layer is made of a number of neurons or nodes. Every neuron in each of these layers is connected to every neuron in the next layer for a network that is fully connected. The number of connections could be less, depending upon the application. Every NN has an input and an output neural layer. There can be some additional layers in between these two, known as the hidden layers.

In general, an NN is very robust: its performance does not degrade even if some of its interconnections are not being used. An NN model is specified by the net topology, node characteristics, and training or learning rules. These rules specify an initial set of weights and indicate how these weights should be adapted during use to improve performance. Such training can be supervised or unsupervised.

The operating principle of an NN can be mathematically represented by a nonlinear function, f , as follows¹:

$$y_i = f \left(\sum_{j=1}^N w_{ij} x_j - \alpha_i \right) , \quad (1)$$

where x_j are N inputs, y_i is the output of i^{th} neuron, w_{ij} are the weights applied by the i^{th} neuron to the input from the j^{th} neuron, and α_i is a threshold term. The nonlinear function used in NN modeling is usually a hard limiter, a sigmoid function, etc.

Much work is going on in the software and hardware implementation of NN's. Both electronic and optical approaches are being used. Optical implementations have a very promising future because of the inherent parallelism of light. Much of this work uses a free-space architecture to implement the vector-matrix multiplication concept embedded in equation (1).⁴⁻⁶ Most of these architectures contain at least three planes: one, corresponding to the input neurons, contains an array of discrete sources of light (either one- or two-dimensional (1-d or 2-d)), i.e., spatial light modulators (SLM's), light-emitting diodes (LED's), etc; the second contains a 2-d array of interconnection elements representing the weight matrix, i.e., an SLM (either fixed or programmable) or a hologram (2-d or volume), etc; and the third, corresponding to the output neurons, contains an array of discrete photodetectors for summing the weighted inputs with a built-in quadratic nonlinearity (sometimes an additional thresholding is provided at this stage using electronics or some other scheme). Some approaches also include hidden layers. Most of these implementations are of the tabletop type where, due to size, mechanical stability is a problem. Recently, this kind of architecture has been implemented in GaAs.⁶

In 1989 the Defense Advanced Research Projects Agency (DARPA) conducted a study on the state of the art in NN research and its future direction.³ It concluded that a marriage of optics and electronics could be the best way to actually implement the NN architectures and that the most promising technology for the implementation of such optoelectronic NN's is based on GaAs and other III-V semiconductor alloys and superlattices. In this paper we present an optical implementation of a feed-forward NN architecture which is being carried out in a monolithic GaAs/AlGaAs waveguide structure, where both the weighting and thresholding are provided by the room-temperature macroscopic nonlinear behavior of an embedded superlattice by application of small electric voltages.

3. Optoelectronic Implementations of NN's

In a hardware implementation of an NN, the most important functional components are discrete sources of light, interconnections, and weighting, summing, thresholding, and output signals. GaAs-based integrated optics offers components that can perform these functions: laser diodes, low-loss single-mode rib waveguides, power combiners/splitters, modulators, thresholding devices, detectors, etc. A waveguide structure is similar to an optical fiber. The propagating light mode is confined in a very thin region, the core, which is surrounded by layers of material with lower refractive indices known as cladding.

In our waveguide-based architecture of an NN, we have first implemented a single neuron. The sources of light are discrete laser diodes. The light is coupled into

an N -to-1 rib waveguide power combiner structure. The simplest of such structures is a 2-to-1 power combiner, which can be a small-angle Y-junction. The device architecture contains at least three electrically active components (i.e., two modulators and a thresholding element), which are electrically isolated. Ideally, the waveguide parameters should be chosen so that the light propagating through the rib waveguide is single-mode in both transverse and lateral directions. The transverse mode confinement is due to the waveguide core and the lateral confinement is provided by the rib walls. Inclusion of a superlattice in the waveguide core gives rise to electroabsorption. As each of the input signals passes through the device, an electric field is applied to change the transmission of the core material, and the light is modulated: i.e., a synaptic weight is applied. The output guide of the power combiner sums the weighted signals coming from various input guides. It also contains a nonlinear switch in which the waveguide section acts both as a detector and a modulator.^{7,8} The nonlinear switch output depends on the amount of light passing through it; for a higher intensity of light, its transmission is very high, while for a low intensity of light, it is very low.

This power combiner is the basic functional building block of our neural architecture. A more complex structure can be made by cascading many such building blocks. This implementation depends upon designing several integrated optical components: low-loss single-mode rib waveguides, Y-junctions, crossing waveguides with various angles of crossing, modulators, and nonlinear switches. During the course of this investigation, such components have been designed, tested, and integrated monolithically to characterize the performance of a single 2-to-1 neuron.

4. Properties of Superlattices

Advances in III-V optoelectronics research have been particularly rapid in the area of semiconductor laser diodes, detectors, modulators, amplifiers and switches. The ability to grow thin (<10 nm) high-quality layers of compound semiconductors using molecular-beam epitaxy (MBE) has led to the discovery of several new optical and electrical phenomena. Novel optoelectronic devices that depend upon quantum mechanical phenomena have been demonstrated. A quantum well (QW) is a thin heterostructure of a low-bandgap material, such as GaAs, sandwiched between two layers of a higher bandgap material, such as AlGaAs, to constitute an electron well. The thickness of the low-bandgap material is around 30 atomic layers or ~ 10 nm. Large changes in optical absorption due to single and multiple QW's in the presence of an external electrical field have given rise to new optical modulators and switching devices.^{7,8} More recently, attention has been focussed on a periodic structure of very thin QW's (width ~ 3 nm), which are separated by layers of similar width of higher bandgap material. Such a structure of coupled QW's is called a superlattice (SL).

These SL's show very strong macroscopic nonlinear effects at room temperature due to quantum mechanical effects. In such a structure the electrons and holes are not localized, i.e., the energy band structure is continuous. When an electric field is applied, this continuous band structure breaks up into a series of minibands known as the Stark ladder.^{9,10} The separation between these levels depends directly on the applied field, the period of the SL, and the electronic charge. When such a structure is illuminated by light, the absorption of light corresponding to these miniband transitions takes place, as shown in figure 2. This makes the SL absorb light at wavelengths to which it was transparent in the absence of the applied field. Below-gap voltage-modulated absorption devices can also be made when these transitions are made to coincide with the incident photon energy. A great deal of materials engineering is possible with these structures. For example, one can change the layer thickness or the potential well depths to create devices with very different operating characteristics. Very-high-dynamic range modulators are being designed and tested using such SL's.^{11,12}

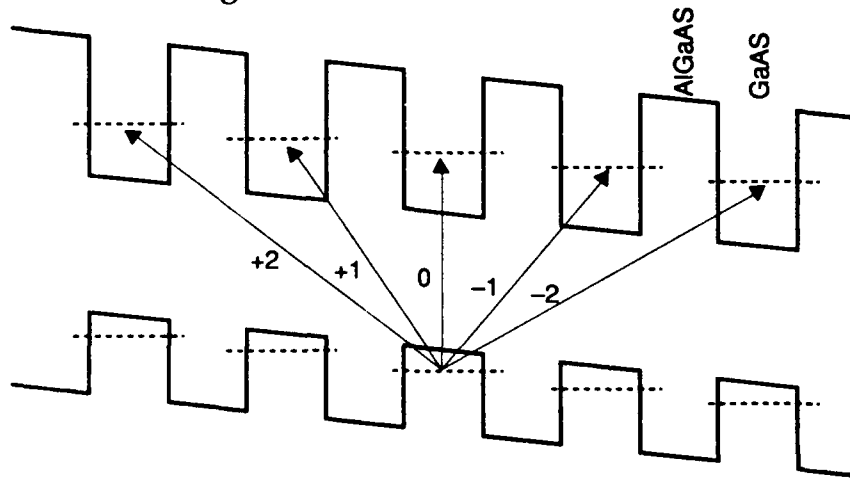
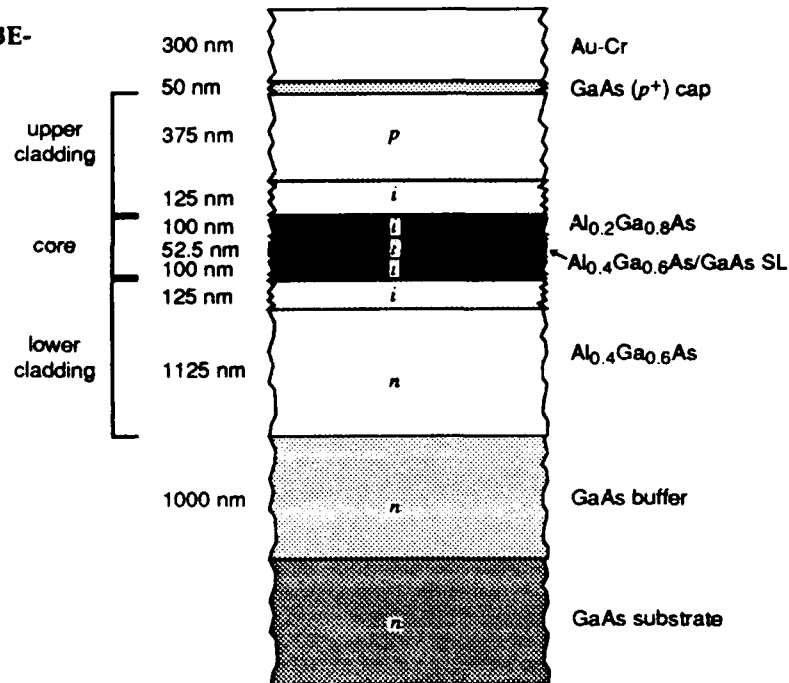


Figure 2. Below-gap transitions for a Stark ladder, giving rise to electroabsorption in presence of an applied field.

5. Material Growth

The GaAs/AlGaAs structure used in this research was grown by MBE. It is grown in a *p-i-n* heterostructure waveguide configuration, as shown in figure 3. The waveguide core is undoped; it consists of an SL 52.5 nm thick, with 100 nm of $\text{Al}_{0.2}\text{Ga}_{0.8}\text{As}$ ($n = 3.53$) on either side. The 10-period SL contains alternate layers of 2.5-nm GaAs and 2.5-nm $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$. The top cladding region contains 500 nm of $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$ ($n = 3.39$), out of which the first 125 nm next to the core are undoped and the rest is *p*-doped. The bottom cladding region also consists of 1250 nm of $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$, including 125 nm of intrinsic layer next to the core, and the rest is *n*-doped. The doping levels are around $10^{18}/\text{cm}^3$. These layers are grown on *n*⁺-doped

Figure 3. Layer structure of MBE-grown material (not drawn to scale).



GaAs substrate with an n -doped 1000-nm GaAs buffer layer. The topmost layer in this structure is the 50-nm p^+ GaAs cap. Incorporation of the p - i - n diode structure is very useful in applying an external electric field to the waveguide core. The doping is spaced back from the core region to reduce free-carrier-induced losses for the propagating optical mode. The thickness of the core region is chosen to give a single transverse electric (TE) or transverse magnetic (TM) propagation mode at a 780-nm wavelength in the transverse direction, as shown in figure 4. From photoluminescence measurements, the bandgap of this material is found to be 1.75 eV or 708 nm.

6. Device Fabrication

A self-aligned single-mask lift-off technique was used to fabricate 5- μm rib waveguides in a Y-junction configuration. This technique consists of using photolithography to pattern the resist mask on the MBE-grown material, depositing 300 nm of chrome/gold film by e -beam evaporation through the developed resist, lifting off the undesired chrome/gold by developing the resist, and then etching the rib pattern by using a standard GaAs chemical etchant, with the metallization providing the electrode and the self-aligned mask. Figure 5(a) shows the device dimensions for a single device, figure 5(b) shows a 3-d illustration of the Y-junction optical neuron, and figure 6 shows a microscope photograph (magnification 24 \times) of the chip containing 10 neuron devices. Each device is 4 mm long; the two input guides are sepa-

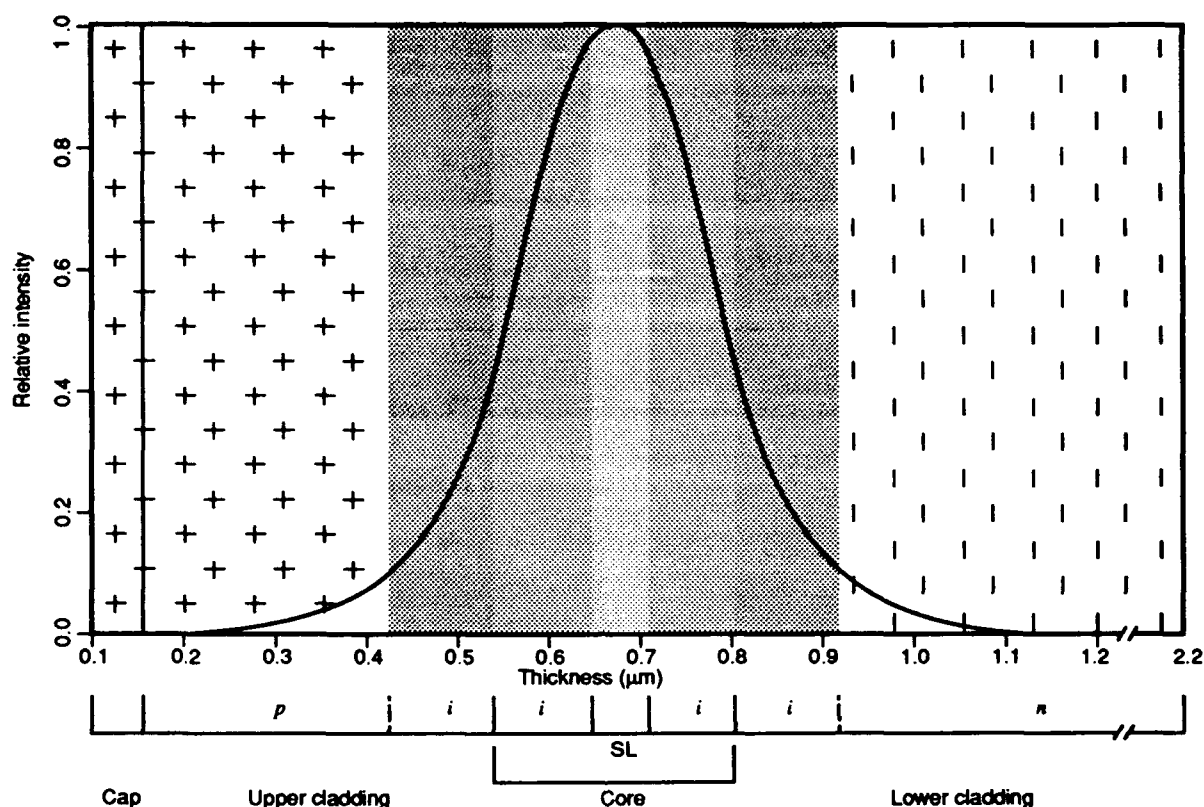


Figure 4. Profile of single transverse mode in waveguide core, showing how intensity of light is distributed as a function of device thickness. Core and cladding regions are shown with their doping profile.

rated by $50\text{ }\mu\text{m}$. The height of the rib walls was chosen to be 550 nm , which is enough for electrical isolation of the device but makes the device multimode in the lateral direction, thus degrading the SNR slightly. Two cuts ($3 \times 5\text{ }\mu\text{m}$) were included in the Y-junction mask, which resulted during etching produced the breaks needed to electrically isolate the three active components, labelled M1, M2, and M3, as shown in figure 5(b). The entire chip is cleaved at both ends to provide optical-quality coupling facets. The chip is mounted on a conducting holder which acts as the bottom n -contact, and p -contacts are provided by the top gold layer. Three $100 \times 200\text{ }\mu\text{m}$ bonding pads connected by $10\text{-}\mu\text{m}$ -wide conducting channels were also incorporated in the device. The chip is mounted on a holder with pins. Gold wires $25\text{ }\mu\text{m}$ thick are bonded to the p - and n -contacts so that these can be connected to pins for easy electrical access.

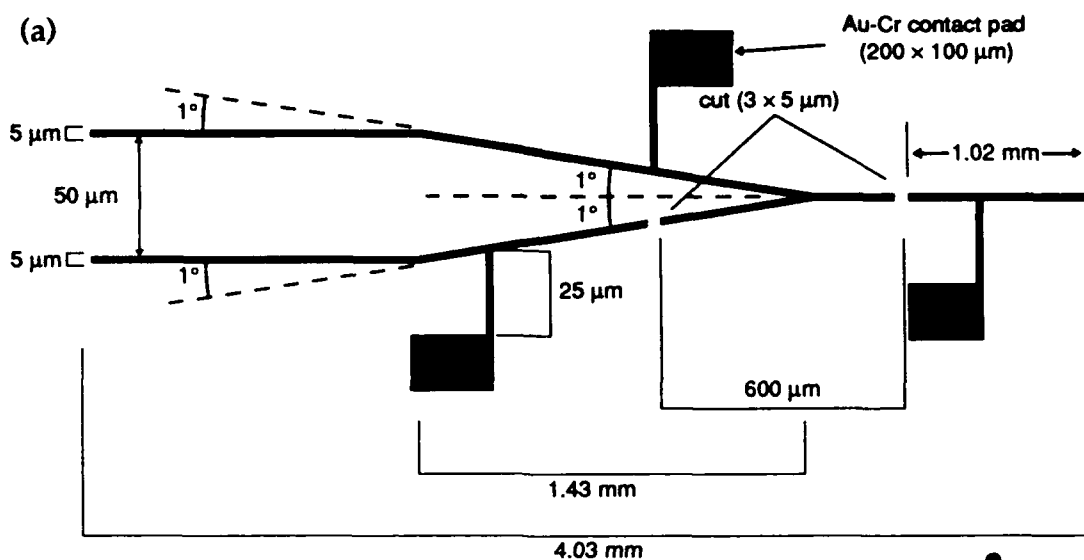


Figure 5. 2-to-1 waveguide neuron: (a) Geometric layout showing dimensions of various parts (not drawn to scale). (b) Schematic 3-d illustration of functional building block; rib waveguide structure, three electrically isolated active components (M1, M2, and M3), and electrical contact pads are clearly shown.

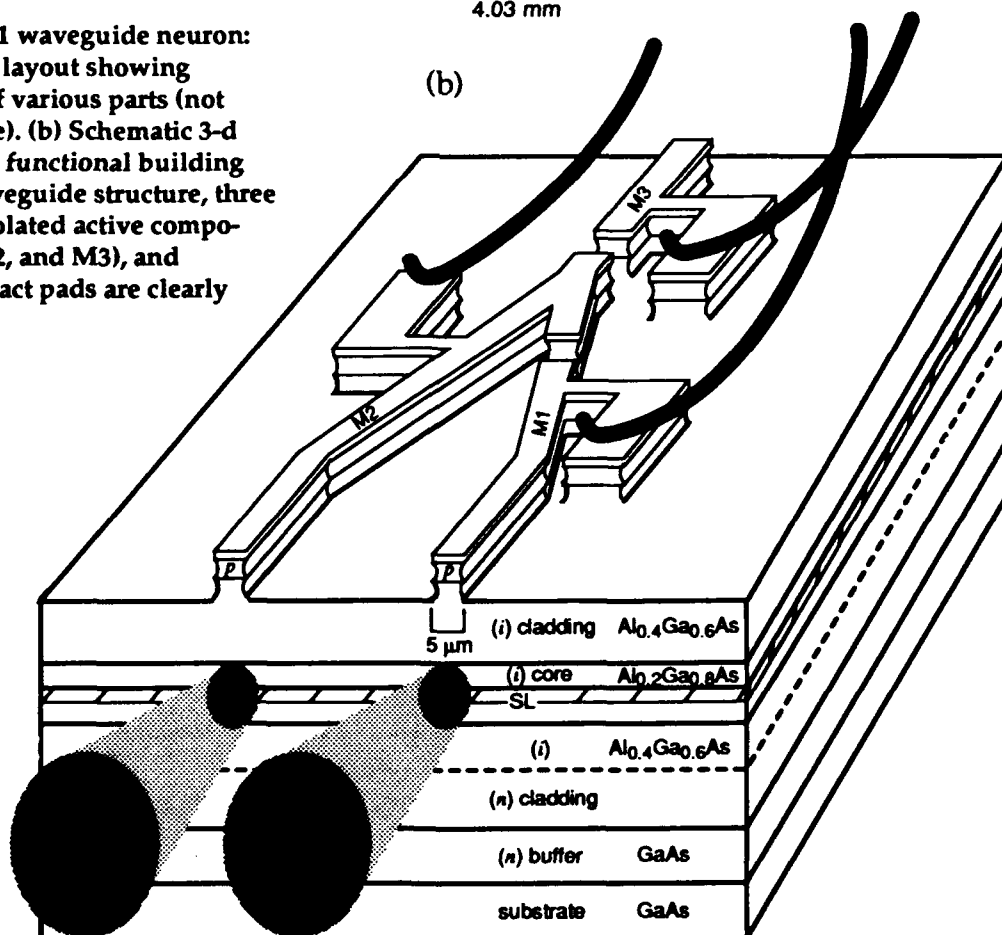
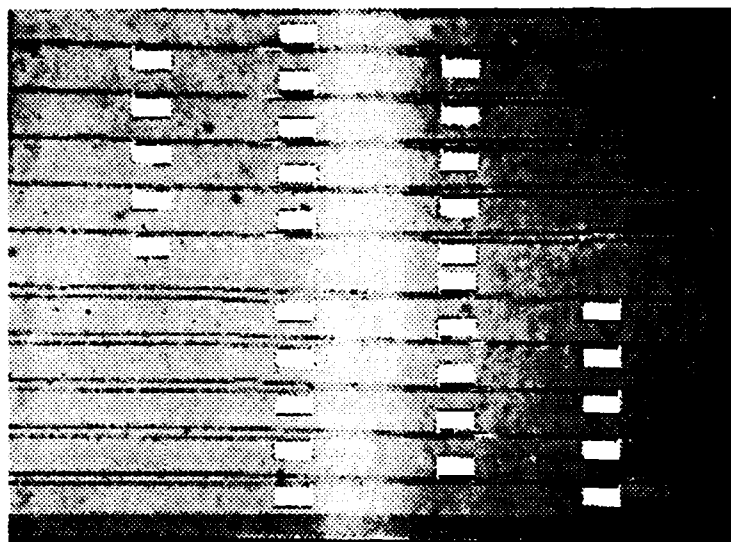


Figure 6. An optical microscope image of 10 waveguide neurons with a magnification of 24 \times . First five neurons are arranged from right to left, rest are in opposite direction. Each neuron device has dimensions given in figure 5(a).



7. Experimental Setup

The experimental setup is shown in figures 7(a) and (b). Light from two 780-nm diode lasers (Sharp LT024) is end-fire coupled into the two input guides of the device using a 40 \times microscope objective to focus the two incoming beams of light. We separate the two focussed spots by 50 μm by choosing suitable position alignments for the two lasers. The two laser beams used are totally incoherent with each other. Only TE-polarized light is coupled into the waveguide. The output of the waveguide device is collected by a 10 \times microscope objective and focussed onto a spatial filter. The light is next focussed by a lens onto a detector connected to a power meter and a charge-coupled device (CCD) camera. The image is processed using a PC with a frame-grabber. The spatial filter is used to block out light not coming from the waveguide. Another CCD camera along with a microscope objective is used for the top viewing of the sample and is very useful in positioning and alignment of the laser inputs. The electrical voltages are applied to the device using a four-wire cable, which is connected to the contact pins on the sample holder. The electronic bias circuit is designed to limit the current flowing in each of the three components to 70 μA . Since the positioning requirements for the optical coupling of the laser beams are very demanding (the focussed spot has to be coupled into a $5 \times 0.25 \mu\text{m}$ region on the input facet), a piezoelectric transducer (PZT) three-axis controller is used for fine position alignment of the sample.

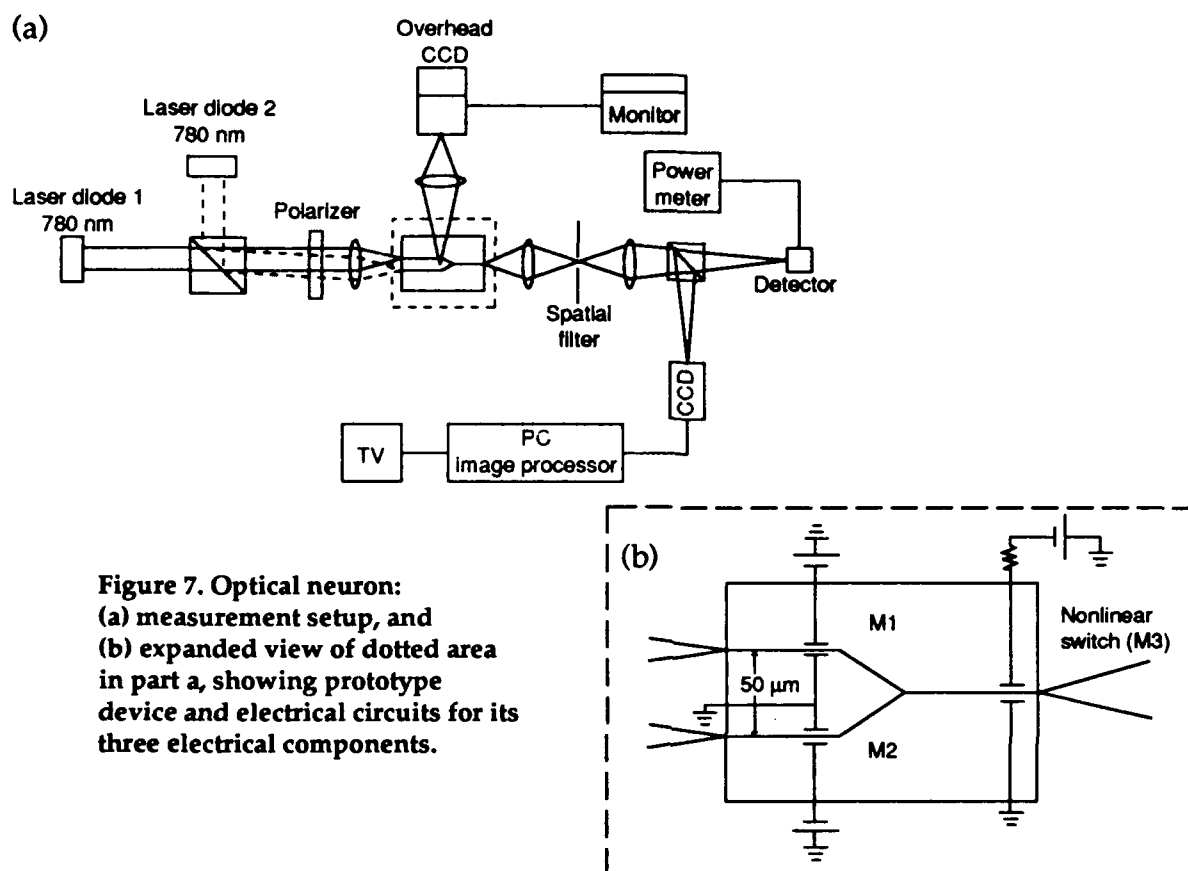


Figure 7. Optical neuron:
 (a) measurement setup, and
 (b) expanded view of dotted area
 in part a, showing prototype
 device and electrical circuits for its
 three electrical components.

8. Results

The three active components on the device, M1, M2, and M3 (see fig. 5(b)), were first characterized as modulators. Figure 8 shows the results for M3, which is 1 mm long. The output intensity of the device is measured as a function of the applied bias. The device transmits less light as the reverse-bias voltage is increased, because of the absorptions caused by the miniband transitions. Transmission modulation as large as 250 dB/cm has been observed for TE polarized light at 780 nm. M1 and M2 perform similarly. Similar results have also been obtained for a range of wavelengths from 777 to 810 nm using a titanium-sapphire laser.¹² The transmitted signal out of a modulator corresponds to weighted input signals according to equation (1).

We also tested M3 to perform as an optical thresholding device by using it both as a modulator and a detector.^{7,8} This was done by inserting a large series resistor in the modulator reverse-bias circuit at a fixed voltage. In this configuration, when the incident light is absorbed by the device and a current starts to flow in the circuit,

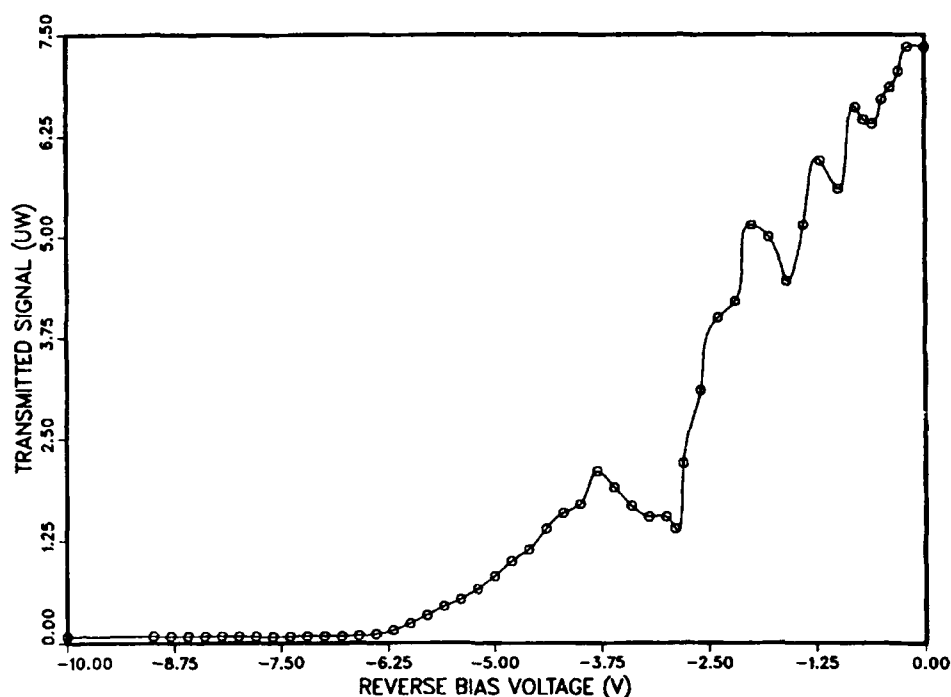


Figure 8. Measurement results showing transmitted output intensity as a function of reverse-bias voltage at 780 nm for TE polarization of incident light. Modulator M3 is 1 mm long.

there is a voltage drop across the resistor which reduces the device voltage, making it transmit more light. Now if the incident light on the device is reduced in intensity, it reduces the current flowing through the circuit (one electron-hole pair is created for every photon absorbed), which reduces the voltage drop across the resistor and increases it across the device, making it transmit less light. Measuring the change in relative transmission of the device as a function of the relative input intensity is a very useful measure of its nonlinear behavior. Figure 9 shows the results obtained under two different reverse bias voltage/resistor settings: one with -10 V and 5 M Ω , and second with -7 V and 1 M Ω . The relative transmission changed from 0 to 0.95 for the first curve and from 0 to 0.87 for the second one. The modulation depth for this switch is around 25 dB. These results clearly show that the nonlinear behavior can be dynamically adjusted based on the system requirements.

Finally, the performance of a 2-to-1 waveguide neuron was evaluated, and the results look very good. The nonlinear switch was run with -10 V and 5 M Ω . The device was trained to test an AND logic in which the two sets of weights were selected for different input intensities so that the device is on when the total input

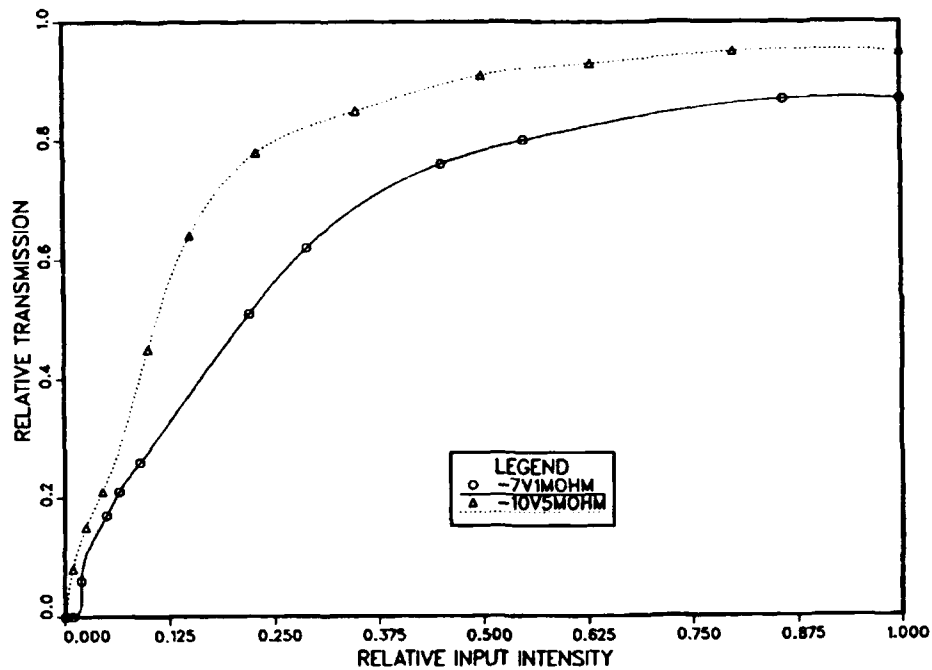


Figure 9. Measurement results showing performance of nonlinear switch. Relative transmission is plotted as a function of relative input intensity for a 1-mm-long component (M3) with -10-V bias and $5\text{-M}\Omega$ series resistor, and with -7-V bias and $1\text{-M}\Omega$ series resistor.

intensity of both lasers ($x_1 + x_2$) is over a threshold value; otherwise it was off. The results are listed in table 1.

The device was run with four different sets of operating conditions. In the first set, the output of the device for a particular input intensity was measured to be $2.4\text{ }\mu\text{W}$ without any of the active components on. Second, M3 was turned on with -10 V and $5\text{ M}\Omega$, and the output was $2.3\text{ }\mu\text{W}$ for the same conditions. Third, the input was dropped by a factor of seven (the threshold value), and the output was then 230 nW . Fourth, M1 was turned on with -4.25 V , and the output was 1 nW . The neuron on/off ratio of 2300 was obtained by comparing set 4 with set 2 and it is mostly input signal limited.

Table 1. Performance of 2-to-1 optical neuron

Input set	M1 (V)	M2 (V)	M3	Output intensity	Neuron signal (y)
1	0	0	0 V	$2.4\text{ }\mu\text{W}$	On
2	0	0	$-10\text{ V}, 5\text{ M}\Omega$	$2.3\text{ }\mu\text{W}$	On
3	0	0	$-10\text{ V}, 5\text{ M}\Omega$	230 nW	On
4	-4.25	0	$-10\text{ V}, 5\text{ M}\Omega$	1 nW	Off

As discussed earlier, the waveguide implementation of an NN requires designing low-loss single-mode rib waveguides. The propagation losses for the MBE-grown material were measured using the standard cut-back method. In this method, several lengths of planar waveguide samples were used to measure the loss. The measured losses are ~ 43 dB/cm at 780 nm. The rib waveguide losses would be even higher. Additional losses are contributed by the coupling of input signals into the guide and the bends in the structure.

9. Conclusions and Discussion

An optoelectronic implementation of an NN in GaAs is described, and experimental results are presented. A 2-to-1 neuron device was fabricated and characterized for performance. The results look very good. A 250-dB/cm below-gap optical modulator and a 25-dB optoelectronic nonlinear switch have been tested. These high-dynamic-range modulators have many potential applications. The nonlinear device described in this work can be used as a thresholding detector device, and an array of such devices can be used in other optical implementations of an NN architecture. In future, laser diodes may be integrated on the chip. The single neuron in this architecture has multiple inputs, unlike many other implementations.

Several problems were encountered in the development. The two most important ones are the high propagation loss of the material, and the packing density of the neurons due to the relatively large area ($4\text{ mm} \times 200\text{ }\mu\text{m}$) required in the fabrication of each neuron. Because small-angle Y-junctions are used, the waveguide needs to be relatively long (1.43 mm) to provide $50\text{-}\mu\text{m}$ separation between the two input guides (see fig. 5). Each modulator also needs to be relatively long for this reason, as well as for adequate modulation. These problems have an impact on the future design of more complex single neurons and an NN.

The first problem can be addressed by growing the p -doped region in the upper cladding much farther back from the core region. We have already accomplished this by growing new MBE material with propagation losses of <10 dB/cm. With this material it will be possible to design a single lateral-mode rib waveguide to give better SNR.

The approach to solving the next problem is twofold: First, material is grown in which the core region is made up of 100-percent SL instead of only 20-percent SL, as discussed in section 5. Such a material has already been grown. This may increase the modulation and thresholding by as much as a factor of five, which means that the total device length can be reduced correspondingly. The second approach is to design a new compact type of power combiner that uses either S-bends or right-

angled bends.^{13,14} This work is also proceeding. New approaches for electrical contacts are also being explored.

After a 2-to-1 neuron has been optimized and characterized, more complex neurons (4-to-1) and a small NN will be designed and tested.

Acknowledgement

We thank Fred Towner of Martin Marietta Laboratories, Baltimore, for his assistance in the MBE sample growth.

References

1. R. P. Lippmann, An introduction to computing with neural nets, *IEEE ASSP Magazine* 4 (April 1987).
2. B. Widrow, R. G. Winter, and R. J. Baxter, Layered neural nets for pattern recognition, *IEEE Trans. Acoust. Speech Signal Process.* 36, No. 7, 1109 (1988).
3. *DARPA Neural Network Study*, AECEA Int. Press, Fairfax, VA (November 1988).
4. N. H. Farhat, Optoelectronic neural networks and learning machine, *IEEE Circuits and Devices Magazine* 32 (September 1989).
5. D. Psaltis, D. Brady, X. Gu, and K. Hsu, Optical implementation of neural computers, *Optical Processing and Computing*, ed. H. A. Arsenault, ch 8 (1989).
6. Y. Nitta, J. Ohita, M. Takahashi, S. Tai, and K. Kyuma, Optical neurochip with learning capability, *Photonics Tech. Lett.* 4, No. 3, 247 (1992).
7. D.A.B. Miller, Optoelectronic applications of quantum wells, *Opt. Photonics News* 7 (February 1990).
8. D.A.B. Miller, Quantum-well self-electro-optic effect devices, *Opt. Quantum Electron.* 22, S61 (1990).
9. R. P. Leavitt and J. W. Little, Stark ladders in strongly coupled superlattices and their interactions with embedded quantum wells, *Phys. Rev.* B41, 5174 (1990).
10. R. P. Leavitt and J. W. Little, Excitonic effects in the optical spectra of superlattices in an electric field, *Phys. Rev.* B42, 11,784 (1990).
11. E. Bigan, M. Allovon, M. Carre, C. Braud, A. Carencu, and P. Voisin, Optimization of optical waveguide modulators based on Wannier-Stark localization: An experimental study, *IEEE J. Quantum Electron.* 28, 214 (1992).
12. M. R. Stead, G. Simonis, J. Pham, R. P. Leavitt, N. Gupta, and J. Bradshaw, Characterization of Stark-Wannier waveguide modulator, paper to be presented at CLEO '92.
13. C. Rolland, D. M. Adams, and D. Yevick, Optimization of strongly guiding semiconductor rib waveguide Y-junction, *IEEE Photonic Tech. Lett.* 2, 404 (1990).
14. C. T. Sullivan, *III-V Optical Waveguide Technology*, Honeywell Systems and Research Center, Air Force WRDC-TR-90-5027 (January 1991).